



# TRANSLATOR CERTIFICATION

450 7th Ave | 6th Floor | New York, NY 10123 | Tel: 212.643.8800 | Fax: 212.643.0005 | [www.mside.com](http://www.mside.com)

Morningside | Translations

I, Shoko Sato, a translator fluent in the Japanese and English languages, on behalf of Morningside Evaluations and Consulting, do solemnly and sincerely declare that the following is, to the best of my knowledge and belief, a true and correct translation of the document(s) listed below in a form that best reflects the intention and meaning of the original text.

## MORNINGSIDE EVALUATIONS AND CONSULTING

\_\_\_\_\_  
Signature of Translator

Date: May 15, 2008

Description of Documents Translated:

Japanese Patent Application WO 02/0881792

---



(19) International Bureau of World  
Intellectual Property Organization (WIPO)



(43) International Publication Date:  
November 7, 2002

PCT

(10) International Laid-Open No.

WO 02/088179 A1

(51) International Patent Classification<sup>7</sup>: C07K 14/415, C12N  
15/29, A61H 1/00, S18, C12Q 1/68

(72) Inventor; and

(75) Inventor/ applicant (only for US): Jun Imamura [JP/JP]; c/o  
Plantech Research Institute, 1000, Kamoshida-cho, Aoba-ku,  
Yokohama-shi, Kanagawa-ken 227-0033 (JP); Hideya Fujimoto  
[JP/JP]; c/o Plantech Research Institute, 1000, Kamoshida-cho,  
Aoba-ku, Yokohama-shi, Kanagawa-ken 227-0033 (JP); Ritsuko  
Imai [JP/JP]; c/o Plantech Research Institute, 1000, Kamoshida-  
cho, Aoba-ku, Yokohama-shi, Kanagawa-ken 227-0033 (JP);  
Nobuya Koizuka [JP/JP]; c/o Plantech Research Institute, 1000,  
Kamoshida-cho, Aoba-ku, Yokohama-shi, Kanagawa-ken 227-  
0033 (JP); Takako Sakai [JP/JP]; c/o Plantech Research  
Institute, 1000, Kamoshida-cho, Aoba-ku, Yokohama-shi,  
Kanagawa-ken 227-0033 (JP); Takahiko Hayakawa [JP/JP]; c/o  
Plantech Research Institute, 1000, Kamoshida-cho, Aoba-ku,  
Yokohama-shi, Kanagawa-ken 227-0033 (JP)

(21) International Application Number: PCT/JP02/04092

(22) International Filing Date: April 24, 2002

(25) Language of International Application: Japanese

(26) Language of International Publication: Japanese

(30) Priority data:

Japanese Patent Application 2001-128008 April 25, 2001 JP  
Japanese Patent Application 2001-202082 July 3, 2001 JP  
Japanese Patent Application 2002-20083 March 29, 2002 JP

(71) Applicant (for all designated states except for US): Mitsubishi  
Chemical Corporation) [JP/JP]; 5-2, Marunouchi 2-chome,  
Chiyoda-ku, Tokyo 100-0005 (JP)

(74) Agent: Patent Corporation Patent Office SIKS & CO.; Kyobashi  
Nishoku Bldg. 8F, 8-7, Kyobashi 1-chome, Chuo-ku, Tokyo  
104-0031 (JP)

[Continued]

(54) Title: PROTEIN PARTICIPATING IN RESTORATION FROM CYTOPLASMIC MALE STERILITY TO FERTILITY AND GENE ENCODING THE SAME

(54) PROTEINS PARTICIPATING IN RESTORATION FROM CYTOPLASMIC MALE STERILITY TO FERTILITY AND GENES CODING THEREFOR

(57) Abstract: It is intended to isolate an Rf gene, in particular a radish-origin Rf1 gene, and identify its structure. A protein participating in restoration of cytoplasmic male sterility to fertility characterized by having at least 14 pentatricopeptide repeat (PPR) motifs wherein these motifs are divided into at least 3 groups each having at least 2 PPR motifs and the block in the carboxy-terminal (C-terminal) side has 4 PPR motifs.

(57) [Abstract]

The objective of the present invention is to isolate Rf genes, particularly, a Rf1 gene derived from radish and to identify the structure thereof. According to the present invention, the proteins contain 14 or more of pentatricopeptide repeat (hereafter, abbreviated as PPR) motifs: the PPR motif group is divided into three or more blocks; each block has at least two or more PPR motifs, respectively; and the block at a carboxy terminal (C-terminal) has four PPR motifs



Designated states  
(81) (domestic): AF, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, IT, KG, KR, KZ, LC, LK, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

Designated states  
(84) (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasia patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent

(AE, BF, CH, CY, DE, DK, FS, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NF, SN, TD, TG).

Attached publication document:  
International Search Report

For two-letter codes and other abbreviations, see "Guidance note for codes and abbreviations" published on the first page of the text of each of regularly issued PCT Gazettes.

## Specification

PROTEINS PARTICIPATING IN RESTORATION FROM CYTOPLASMIC  
MALE STERILITY TO FERTILITY AND GENES CODING THEREFOR

## Technical Field

The present invention relates to genes participating in restoration of cytoplasmic male sterility to fertility. In further detail, the present invention relates to genes participating in the restoration of cytoplasmic male sterility (hereafter, it may be abbreviated as 'cms') character to be utilized for the purpose of F1 hybrid (hereafter, referred to as F1) breed development; a vector containing said genes; and a transformant.

## Background Technology

In cereal crops, such as grain crops or vegetables, F1 breed development has caught on based upon the characteristics including 1) superior agronomic character due to hybrid vigor, 2) uniformity of harvest, and 3) protection of profits by breed cultivators because the genetic character is separated in the next generation, and this is practically used in many main crops.

As one of methods for producing F1 variety seeds, a cms/Rf seed production system comprising the cytoplasmic male sterility (cms) line and a line for restoring the male sterility (hereafter, this may be abbreviated as 'Rf') is available, and for example, this system is developed with cereal crops, such as rice, sorghum or corn, and oil crops, such as sunflower; however, these are all developed using a technique, such as hybridization or cell fusion.

In the meantime, F1 seed production system using self-incompatibility has been widely utilized in cruciferous; however, regarding rapeseed, because it has no stable self-incompatibility, another F1 seed production system utilizing the cms line and the Rf line is in demand.

In the meantime, a study where the cytoplasmic male sterility (kosena-cms) derived from kosena radish (*Raphanus sativus* cv. Kosena) and the cytoplasmic male sterility (ogura-cms) derived from ogura radish (*Raphanus sativus* L.) are used in rapeseed has been continued in recent years. Both cms genes are coded to the genome of mitochondria, which are cytoplasmic organelle, and their base sequences are also known; however, for radish, since molecular-biological research is not in progress and markers that are required for gene isolation are hardly known, as well, it is difficult to isolate genes from the nuclear, and regarding Rf, it is merely introduced to rapeseed using techniques of hybridization or cell fusion from a fertility restoration system for radish. In addition, regarding the Rf gene, one or multiple restoring genes depending upon each cms system of plants. In the radish, it is necessary for the Rf1 and Rf2 genes simultaneously exist for the fertility restoration, and in addition, it is known that the Rf1 gene drastically decreases accumulation of ORF125 protein (M. Iwabuchi et al. Plant Mol. Biol. 39:183-188, 1999) within the mitochondria known as a cms causal protein of radish (Japanese Journal of Breeding 47 (supplementary volume 1) p. 186, 1997, Japanese Journal of Breeding 48 (supplementary volume 1) p. 197, 1998).

Further, in the rapeseed, it is known that the Rf1 gene of radish introduced by hybridization or cell fusion according to a gene analysis test decreases the accumulation of ORF125 or ORF138 protein (M. Grelon et al. Mol. Gen. Genet. 243:540-547), which is known as a cms causal protein, and the decrease in the accumulation and the phenomenon of restoring the fertility in this ORF125 or ORF138 protein are completely matched (N. Koizuka, et al. Theor Appl Genet, 100:949-955, 2000). In other words, in order to restore the fertility with the rapeseed male sterile line, it is essential to decrease the accumulation of ORF125 or ORF138 protein, and the Rf1 gene is the important gene for this purpose.

However, for base sequence of the Rf gene, although only the Rf2 gene, which is one of restoring genes for T-cytoplasm that is one of corn cms, is identified and isolated; however, any base sequence of Rf genes in other plants is not known at all.

In the rapeseed restoring line where Rf1 gene derived from the Ogura radish (*Raphanus sativus* L.) has been introduced by hybridization or cell fusion and the F1 variety that has been produced by said line as a father, the glucosinolate (hereafter, abbreviated as GSL) content becomes higher than a regulation value, and this is a practical problem. It appears that because the radish-derived gene participating in the GSL biosynthesis exists in the vicinity of the Rf1 gene and is genetically linked strongly, the GSL content of the rapeseed restoring line (Rf line) is increased. GSL is contained in oil expression, and since it is known that when GSL is supplied to animals as feed, it causes enlarged thyroid gland, the GSL content of rapeseed is required to be 18  $\mu$ mole/g in North America and 20  $\mu$ mole/g or less in Europe during development of new varieties plants.

## DISCLOSURE OF THE INVENTION

In a rapeseed restoration line, where a Rf1 gene derived from ogura radish has been introduced due to hybridization or cell fusion, and F1 variety produced using the line as a father; it has been ascertained that a glucosinolate content (hereafter, abbreviated as GSL) becomes higher than a regulation value, and this is a practical problem. It is believed that this is because the radish-derived genes participating in GSL biosynthesis exist in the vicinity of Rf1 gene and are genetically strongly linked, and the GSL content of the rapeseed restoration line (Rf line) is increased. GSL is contained in rape cakes, and since it is known that when it is supplied to animals as feed, this causes hypertrophied thyroid gland, the GSL content of

rapeseed requires being 18  $\mu\text{mole/g}$  in North America and 20  $\mu\text{mole/g}$  or less in Europe during a breeding stage.

In addition, in recent years, the development of plants where a function, such as herbicide resistance, is added due to genetic recombination is on the move, and in order to efficiently create these plants, only the presence of rapeseed restoration line obtained by hybridization or cell fusion is not sufficient, and the isolation of Rf gene, in particular, radish-derived Rf1 gene has been in demand.

In other words, isolation of the Rf gene, in particular, the radish-derived Rf1 gene, and identification of the structure thereof were designated as a problem to be solved. In addition, another problem to be solved in the present invention is to provide a means for establishing the rapeseed restoration line by utilizing the isolated Rf gene.

As a result of a keen examination for solving the problems, the inventors of the present invention have succeeded to clone the Rf1 genes from radish and rapeseed, and to solve the problems.

In other words, according to the present invention, proteins participating in restoration from cytoplasmic male sterility to fertility, wherein the proteins contain 14 or more of pentatricopeptide repeat (hereafter, abbreviated as PPR) motifs; the PPR motif group is divided into three or more blocks; each block has at least two or more PPR motifs, respectively; and the block at a carboxy terminal (C-terminal) has four PPR motifs are provided.

According to preferred embodiments of the proteins,  
proteins whose number of PPR motifs is 14 to 16;  
proteins where the PPR motif group is divided into three blocks, and the blocks have five, seven and four PPR motifs in respective order from the amino terminal (N-terminal) side, respectively;

proteins where the fourth amino acid existing in the second PPR motif existing from the amino terminal (N-terminal) side is other than serine, threonine and cysteine;

proteins where the fourth amino acid existing in the second PPR motif from the amino terminal (N-terminal) is any of asparagine, glutamine, asparagine acid, glutamic acid and histidine; and

proteins further containing a signal peptide sequence for migrating to mitochondria at the amino terminal, or a sequence comprising - LysAspGluLeu- at the carboxy terminal; are provided.

According to another aspect of the present invention, proteins participating in restoration from cytoplasmic male sterility to fertility causing an occurrence of a gel shift to the transcription after making contact with a transcription of a cytoplasmic male sterile gene.

According to further another aspect of the present invention, a protein participating in the restoration of cytoplasmic male sterile individual having amino acid sequence described in SEQ ID NO 26 from sterility to fertility;

a protein participating in the restoration of cytoplasmic male sterile individual having amino acid sequence described in SEQ ID NO 27 from sterility to fertility;

a protein participating in the restoration of cytoplasmic male sterile individual having amino acid sequence described in SEQ ID NO 28 from sterility to fertility;

a protein participating in the restoration of cytoplasmic male sterile individual having amino acid sequence described in SEQ ID NO 29 from sterility to fertility; are provided.

According to another aspect of the present invention, either of the following proteins is provided:

(1) a protein having the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 3, the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 17, and the 82nd to 690th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 19; or

(2) a protein having amino-acid sequences where one to multiple amino acids are deleted, added and/or substituted in the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 3, the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 17 and the 82nd to 690th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 19 and participating in restoration from cytoplasmic male sterility to fertility.

According to another aspect of the present invention, either of the following proteins is provided:

(1) a protein having amino-acid sequences described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19; or

(2) a protein having an amino-acid sequence where one to multiple amino acids are deleted, added and/or substituted described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19 and participating in restoration from cytoplasmic male sterility to fertility.

In the present invention, preferably, the cytoplasmic male sterile individual has a cytoplasmic male sterile gene of kosena radish (*Raphanus sativus* cv. Kosena) and/or Ogura radish (*Raphanus sativus* L.) or homologue thereof.

In the present invention, preferably, the cytoplasmic male sterile individual has a cytoplasmic male sterile gene(s) of kosena radish (*Raphanus sativus* cv. Kosena) and/or ogura radish (*Raphanus sativus* L.) or homologue thereof.

According to another aspect of the present invention, DNA coding for either of the above-mentioned proteins is provided.

According to another aspect of the present invention,  
DNA having base sequences described in SEQ ID NO 22;  
DNA having base sequences described in SEQ ID NO 23;  
DNA having base sequences described in SEQ ID NO 24; and  
DNA having base sequences described in SEQ ID NO 25;  
are provided.

According to another aspect of the present invention, any of the following DNA is provided:

- (1) DNA having base sequences described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18; or
- (2) DNA having base sequences where one to multiple bases are deleted, added and/or substituted in the base sequences described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18 and participating in restoration from cytoplasmic male sterility to fertility; or
- (3) DNA hybridizing with base sequences described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18 under stringent conditions and participating in restoration from cytoplasmic male sterility to fertility.

According to another aspect of the present invention, any of the following DNA is provided:

- (1) DNA having the 3,754th to 8,553th base sequences among the base sequences described in SEQ ID NO 1 or the 812th to 3002nd base sequences among the base sequences described in SEQ ID NO 15; or
- (2) DNA having the 3,754th to 8,553th base sequences where one to multiple bases are deleted, added and/or substituted among the base sequences described in SEQ ID NO 1 or the 812th to 3002nd base sequences among the base sequences described in SEQ ID NO 15 and participating in restoration from cytoplasmic male sterility to fertility; or
- (3) DNA hybridizing with DNA having the 3,754th to 8,553th base sequences among the base sequences described in SEQ ID NO 1 or the 812th to 3002nd base sequences among the base sequences described in SEQ ID NO 15 under

stringent conditions and participating in restoration from cytoplasmic male sterility to fertility.

According to another aspect of the present invention, any of the following DNA is provided:

(1) DNA having base sequences described in SEQ ID NO 1 or SEQ ID NO 15;

or

(2) DNA having base sequences where one to multiple bases are deleted, added and/or substituted among the base sequences described in SEQ ID NO 1 or SEQ ID NO 15, and participating in restoration from cytoplasmic male sterility to fertility; or

(3) DNA hybridizing with DNA having base sequences described in SEQ ID NO 1 or SEQ ID NO 15 under stringent conditions and participating in restoration from cytoplasmic male sterility to fertility.

According to another aspect of the present invention, the cytoplasmic male sterile individual has a cytoplasmic male sterile gene of kosena radish (*Raphanus sativus* cv. Kosena) and/or ogura radish (*Raphanus sativus* L.) or has homologue thereof.

According to another aspect of the present invention, a vector containing the DNA of the present invention is provided.

According to another aspect of the present invention, a transformant having the DNA of the present invention or the vector of the present invention is provided. The transformant is preferably a transformed plant.

According to another aspect of the present invention, a method for restoring from cytoplasmic male sterility to fertility characterized by using the DNA of the present invention is provided.

According to another aspect of the present invention, a transformant enabling the control of expression of a cytoplasmic male sterile gene by partially or entirely introducing the DNA of the present invention to a cell having a cytoplasmic male sterile gene and having the DNA of the present invention is provided.

According to another aspect of the present invention, a maintenance method for a cytoplasmic male sterile line by utilizing the transformant is provided.

According to another aspect of the present invention, a method for detecting a gene participating in restoration from cytoplasmic male sterility by using 15 mer to 50 mer of oligonucleotide primer arbitrarily set from the DNA of the present invention or at least 15 mer or more of probe comprising all or part of the DNA of the present invention, and by confirming that a quantity of base sequences to be amplified by the primer or a quantity of base sequences detected by the probe is one gene or more per genome in a target biological sample is provided.

According to another aspect of the present invention, a promoter DNA having the 3,754th to 5,091st base sequences among the base sequences described in SEQ ID NO 1 or the 1st to 811th base sequences among the base sequences described in SEQ ID NO 15 is provided.

## BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 shows an Rf marker gene map.

Fig. 2 shows a pattern diagram of lambda clone CHI structure where base sequences described in SEQ ID NO 1 are sufficiently maintained.

Fig. 3 shows detection results of introduced DNA in the transformant using the PCR method.

Lane 1: control vector, Lane 2: transformed rapeseed, Lane 3: cytoplasmic male sterile rapeseed

a: 3,186bp to 3,753bp, length: 568bp

b: 4,869bp to 5,112bp, length: 244bp

c: 7,766bp to 8,250bp, length: 485bp

Fig. 4 shows analysis results of accumulation decrease in ORF125, which is a cms protein in the transformant.

Lane 1: Cytoplasmic male sterile rapeseed -1- 15 µg

Lane 2: Fertility restored rapeseed 15 µg

Lane 3: Cytoplasmic male sterile rapeseed -2- 15 µg

Lanes 4 to 7: Cytoplasmic male sterile rapeseed -2-  
15/2 µg, 15/4 µg, 15/8 µg and 15/16 µg of dilute line

Lane 8: Transformed rapeseed 15 µg

Fig. 5 shows results where anthers were extracted from flower bodies of transformed rapeseed and they were observed by a microscope.

Fig. 6 shows base sequence of pSTV125-5' #LA6 and pSTV125-5' #LA12.

## BEST MODE FOR CARRYING OUT THE INVENTION

Hereafter, embodiments of the present invention will be described in detail.

### (1) Mode of protein of the present invention

The protein of the present invention relates to any proteins from (i) to (v).

(i) Proteins participating in restoration from cytoplasmic male sterility to fertility, wherein the proteins contain 14 or more of pentatricopeptide repeat (hereafter, abbreviated as PPR) motifs; the PPR motif group is divided into three or more blocks; each block has at least two or more PPR motifs, respectively; and the block at a carboxy terminal (C-terminal) has four PPR motifs;

(ii) Proteins participating in restoration from cytoplasmic male sterility to fertility causing an occurrence of a gel shift to the transcription after making contact with a transcription of a cytoplasmic male sterile gene.

(iii) Proteins participating in restoration from cytoplasmic male sterility to fertility having amino acid sequences described any of SEQ ID NOS 26 to 29;

(iv) Any of the following proteins:

- (1) A protein having the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 3, the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 17, or the 82nd to 690th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 19; or
- (2) A protein having an amino-acid sequence where one to multiple amino acids are deleted, added and/or substituted in the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 3, the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 17 or the 82nd to 690th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 19 and participating in restoration from cytoplasmic male sterility to fertility; and
- (v) Any of the following proteins:
- (1) A protein having amino-acid sequences described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19; or
- (2) A protein having an amino-acid sequence where one to multiple amino acids are deleted, added and/or substituted described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19 and participating in restoration from cytoplasmic male sterility to fertility.

In this specification, the PPR motif represents "Pentatricopeptide Repeat" motif. This PPR motif is a motif structure of a new protein discovered as a result of the progress of the genome project of Arabidopsis. The basic motif is such that 35 degenerated amino acid sequences are tandem-repeated on the primary structure of the protein. The PPR motif has a sequence represented by an amino terminal (N-terminal)– "VTYNTLISGYCKNGKLEEALELFEEEMKEKGKIPDV"–carboxy terminal (C-terminal) as a consensus amino acid sequence. This motif is proposed by Small and Peeters (reference: Trends Biochem Sci 2000, 25 46-47), and approximately 200 genes that can adopt this motif in the genome of Arabidopsis are reported in the gene bank, such as GenBank

(<http://www.ncbi.nlm.nih.gov/Genbank/index.html>) at that time when this reference was published. At present, a determination about whether or not one arbitrary protein can have this motif structure can be easily determined by a program in Protein families' database of alignments and HMNs (hereafter, abbreviated as Pfam, <http://www.sanger.ac.uk/Software/Pfam/search.shtml>) located in Sanger Institute, UK.

Examples where a function of the protein having the PPR motif has become clear up to the present date include 1) an example where fermentum PET309, which is a protein to transfer to mitochondria, or *Neurospora crassa* CYA-5 interacts with *cox1* mRNA, which is a mitochondrion gene, and the expression of *cox1* is controlled by processing after transcription or in a translation level (Manthey and McEwen EMBO J 1995 14 4031-4043, Coffin et.al. Curr Genet 1997 32 273-280) and 2) another example where corn CRP1, which is a PPR motif protein to transfer to chloroplast, is essential for translation of *petA* and *petD* genes, which are chloroplast genes, and in addition, is essential for a step of *petD* mRNA processing, as well (Fisk et. al. EMBO J 1999 18 2621-2630); therefore, it is assumed that it is highly possible that the protein having the PPR motif is involved in the translation control in some form.

In this occasion, the inventors of the present application isolated the genes participating in the restoration from the sterility of kosenra radish (*Raphanus sativus* cv. Kosenra) cytoplasmic male sterile individual to fertility, and they discovered that the proteins coded by the genes have 14 or more of pentatricopeptide repeat (hereafter, abbreviated as PPR), and the PPR motif group is divided into three or more blocks, and in addition, each block has at least two or more PPR motifs, and, the block existing at the closest to the carboxy terminal (C-terminal) has four PPR motifs.

As the protein participating in the restoration from the sterility of cytoplasmic male sterile individual to fertility, a protein whose number of

PPR motifs is 14 to 16 is preferable, and a protein where the PPR motif group is divided into three blocks and each block has 5, 7 and 4 PPR motifs in respective order from the amino terminal (N-terminal) side is more preferable.

Specifically, proteins comprises

- (1) PPR cluster #1: a PPR cluster comprising 175 residues where the first PPR motif to the fifth PPR motif from the N-terminal are continuous,
- (2) PPR cluster #2: a PPR cluster comprising 245 residues where the sixth PPR motif to the twelfth PPR motif from the N-terminal are continuous, and
- (3) PPR cluster #3: a PPR cluster comprising 140 residues where the thirteenth PPR motif to the sixteenth PPR motif from the N-terminal are continuous,

The protein where the fourth amino acid existing in the second PPR motif from the amino terminal (N-terminal) is other than serine, threonine or cysteine is further preferable, and the protein where the fourth amino acid existing in the second PPR motif from the amino terminal (N-terminal) side is any of asparagine, glutamine, asparagine acid and histidine is further preferable, and the protein where the fourth amino acid existing in the second PPR motif from the amino terminal (N-terminal) is asparagine is the most preferable.

Since it is normally known that the fertility gene exists in a nuclear genome and the cytoplasmic male sterility exists in mitochondria, it is preferable that the protein participating in the restoration from sterility of the cytoplasmic male sterile individual to fertility has a signal peptide sequence at the amino terminal for transferring to the mitochondria or has a sequence comprising 'LysAspGluLeu' at the carboxy terminal.

The signal peptide at the N-terminal for transferring to the mitochondria includes the one to be confirmed by a prediction program "TargetP" (<http://www.cbs.dtu.dk/services/TargetP/>) based upon the algorithm of O. Emanuelsson et al. (J. Mol. Biol. 300, 1005-1016 (2000)) or the one to be confirmed by another prediction program "Psort" (HYPERLINK

<http://psort.nibb.ac.jp>) <http://psort.nibb.ac.jp/>, and a specific example of the signal peptide includes a signal peptide of Arabidopsis AtOXA1 (W. Sakamoto et al.: Plant Cell Physiol, 41:1157-1163) (MetAlaPheArgGlnThrLeuSerIleArgSerArgLeuPheAlaArgArgAsnGlnProValTyrHisIleIleProArgGluSerAspHisGluArgAsp). Among them, 1 to 79 of amino acid sequences among the amino sequences described in SEQ ID NO 3 are preferable, and 1 to 34 of amino acid sequences among the amino sequences described in SEQ ID NO 3 are especially preferable.

Further, the proteins participating in the restoration from sterility of cytoplasmic male sterile individual to fertility induces a translation inhibition of the cytoplasmic male sterile gene by combining with the transcription of the cytoplasmic male sterile gene, and restores the sterility of the cytoplasmic male sterile individual to fertility.

The transcription of the cytoplasmic male sterile gene includes transcription (mRNA) of genes of ORF125, which is a causal protein causing the kosena cytoplasmic male sterility, or ORF138, which is a causal protein causing the ogura cytoplasmic male sterility, and 5'-UTR (Bonhomme et al: Mol Gen Genet, 235:340-348(1992)) region of the gene is preferable.

A method to confirm about whether or not combining with the transcription of cytoplasmic male sterile gene includes a method where the protein of the present invention is added to mRNA of ORF125 or ORF 138, which was artificially transcribed in vitro, and the mixture is electrophoresed, and a so-called gel shift method, and as specific operation procedures, the gel shift method may be conducted under normal conditions.

Further, another method to confirm whether or not the expression is restrained by adding the protein of the present application to the expression where a fused gene between ORF125 or ORF138 gene and a detectable reporter gene, such as Beta-galactosidase or luciferase, is expressed in Escherichia coli is also included.

Specifically, incorporation of a fertility gene shown with the base sequence of SEQ ID NO 2 into *Escherichia coli* expression vector and another incorporation of a vector where a 5'-UTR region of ORF125 and coding part of 25 amino acids are fused to the LacZ gene into the *Escherichia coli* are prepared, and if the expression is induced, the expression of the LacZ gene is restrained only when the expression vector where the fertility gene is incorporated coexists, and the blue *Escherichia coli* colony becomes white in the presence of X-Gal. As described above, the existence of the function to restore the sterility of the cytoplasmic male sterility to fertility by inducing the translation inhibition of the cytoplasmic male sterile gene can be confirmed via the confirmation by utilizing the gene that codes the protein of the present application.

【 0 0 3 1 】

The most preferable protein among proteins of the present invention participating in the restoration from sterility of cytoplasmic male sterile individual to fertility includes proteins having 70 % or more of homology, preferably 80 % or more, more preferably 90 % or more, further preferably 92 % or more, further more preferably 95 % or more and the most preferably 97 % or more with regard to the amino acid sequences described in SEQ ID NOS 26 to 29, which are consensus sequences. The consensus sequences include the amino acid sequences described in SEQ ID NO 26, preferably the amino acid sequences described in SEQ ID NO 27 or 28, and particularly preferably the amino acid sequences described in SEQ ID NO 29.

Further, specific examples of the proteins include

(1) a protein having the 80th to 687th amino acid sequences among the amino acid sequences described in SEQ ID NO 3, the 80th to 687th amino acid sequences among the amino acid sequences described in SEQ ID NO 17 or the 82nd to 690th amino acid sequences among the amino acid sequences described in SEQ ID NO 19; or

(2) a protein having amino acid sequences where one to multiple amino acids are deleted, added and/or substituted in the 80th to 687th amino acid sequences among the amino acid sequences described in SEQ ID NO 3, the 80th to 687th amino acid sequences among the amino acid sequences described in SEQ ID NO 17 or the 82nd to 690th amino acid sequences among the amino acid sequences described in SEQ ID NO 19 and participating in the restoration from sterility of cytoplasmic male sterile individual to fertility.

Further, the proteins having transfer sequences to mitochondria in the sequences include

(1) a protein having the amino acid sequences described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19; or

(2) a protein having amino acid sequences where one to multiple amino acids are deleted, added and/or substituted in the amino acid sequences described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19 and participating in the restoration from sterility of cytoplasmic male sterile individual to fertility.

The protein of the present invention is a protein that can participate in the restoration from sterility of cytoplasmic male sterile individual to fertility. More specifically, when a transformed plant (Rf line) where DNA that codes for the protein of the present invention is bred with an individual of the cytoplasmic male sterile line (cms line), F1 seeds where the fertility has been restored can be obtained. As preferable individuals of the cms line, individuals where the cytoplasmic male sterile gene(s) of kosena radish (*Raphanus sativus* cv. Kosena) and/or ogura radish (*Raphanus sativus* L.) is introduced are included.

The protein of the present invention can be screened and isolated by utilizing the gel shift method or can be isolated or synthesized by utilizing DNA of the present invention to be described below. The acquisition method for the protein of the present invention will be described later.

(2) DNA mode of the present invention:

The DNA of the present invention relates to any of DNA from the following (i) to (v):

- (i) DNA coding for the proteins of the present invention;
- (ii) DNA having base sequences described in any of SEQ ID NO 22 to SEQ ID NO 25:
- (iii) any of the following DNA:
  - (1) DNA having base sequences described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18; or
  - (2) DNA having base sequences where one to multiple bases are deleted, added or substituted in the base sequence described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18 and participating in the restoration from sterility of an cytoplasmic male sterility to fertility; or
  - (3) DNA hybridizing with DNA having base sequences described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18, and participating in the restoration from sterility of a cytoplasmic male sterile individual to fertility,
- (iv) any of the following DNA:
  - (1) DNA having the 3,754th to 8,553rd base sequences, in base sequences described in SEQ ID NO 1 or the 812th to 3,002th base sequences, in base sequences described in SEQ ID NO 15; or
  - (2) DNA having base sequences where one to multiple bases are deleted, added or substituted the 3,754th to 8,553rd base sequences, in base sequences described in SEQ ID NO 1 or the 812th to 3,002th base sequences, in base sequences described in SEQ ID NO 15; or
  - (3) DNA hybridizing base sequences having the 3,754th to 8,553rd base sequences described in SEQ ID NO 1 or the 812th to 3,002th base sequences, in base sequences described in SEQ ID NO 15 under stringent conditions, and participating in the restoration from sterility of an cytoplasmic male sterile individual to fertility, or

- (v) any of the following DNA:
- (1) DNA base sequences described in SEQ ID NO 1 or SEQ ID NO 15;  
or
  - (2) DNA having base sequences where one to multiple bases are deleted, added or substituted in base sequences described in SEQ ID NO 1 or SEQ ID NO 15; or
  - (3) DNA hybridizing base sequences described in SEQ ID NO 1 or SEQ ID NO 15 under stringent conditions, and participating in the restoration from sterility of a cytoplasmic male sterile individual to fertility.

【 0 0 3 4 】

In this specification, the DNA of the present invention may be referred to as a gene of the present invention.

The base sequences shown with SEQ ID NO 1 are genome DNA base sequences comprising 8,553 bases; and base sequences shown in SEQ ID NO 2 are code sequences acquired from SEQ ID NO 1. The base sequences shown in SEQ ID NO 3 are amino acid sequences coded by the base sequences shown with SEQ ID NO 2.

The base sequences shown in SEQ ID NO 15 are genome DNA base sequences comprising 3,306 bases; and the base sequences shown in SEQ ID NO 16 are code sequences obtained from SEQ ID NO 15. The base sequences shown in SEQ ID NO 17 are amino acid sequences coded by the base sequences shown with SEQ ID NO 16.

In this specification, “base sequences where one to multiple base sequences are deleted, added and/or substituted” mean, for example, that any of 1 to 20 base sequences, preferably 1 to 15 base sequences, more preferably 1 to 10 base sequences, further preferably 1 to 5 base sequences are deleted, added and/or substituted.

In this specification, “amino acid sequences where one to multiple base sequences are deleted, added and/or substituted” mean, for example, that any

of 1 to 20 amino acids, preferably 1 to 15 amino acids, more preferably 1 to 10 amino acids, further preferably 1 to 5 amino acids are deleted, added and/or substituted.

In this specification, "DNA hybridizing under stringent conditions" means a base sequence of DNA acquired by a colony hybridization method, a plaque hybridization method or a Southern blot hybridization method using DNA as a probe, and for example, after hybridization is conducted at 65 degrees C in presence of 0.7 to 1.0 M NaCl using a filter where a colony-derived or plaque-derived DNA or fragment of DNA is immobilized, DNA that can be identified by cleaning a filter under a condition at 65 degrees C using 0.1 to 2 × SSC solution (composition of 1 × SSC is 150 mM sodium chloride and 15 mM sodium citrate).

The hybridization can be conducted based upon a method described in *Molecular Cloning: A laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989 (hereafter, abbreviated as "Molecular Cloning 2nd Ed.).

DNA hybridizing under stringent conditions includes DNA having a certain level or more homology with the base sequence of DNA used as a probe. The 'certain level or more homology' herein is for example, 70 % or more, preferably 80 % or more, more preferably 90 % or more, further preferably 93 % or more and the particularly preferably 95 % or more and the most preferably 97 % or more. Furthermore, the DNA having a certain level or more of homology herein contains both polynucleotide having the above-mentioned homology and complementary strand polynucleotide thereof.

DNA of the present invention is DNA that can participate in the restoration from sterility of cytoplasmic male sterile individual to fertility. More specifically, F1 seeds where fertility is restored can be obtained by breeding a transformed plant (Rf line) where DNA of the present invention is introduced with individuals of cytoplasmic male sterile line (cms line). The individuals of the cms line preferably include an individual where a

cytoplasmic male sterile gene of kosen radish (*Raphanus sativus* cv. Kosen) and/or ogura radish (*Raphanus sativus* L.) is introduced.

(3) Acquisition method for DNA of the present invention:

The acquisition method for DNA of the present invention is not particularly limited. The DNA of the present invention can be isolated or synthesized by utilizing a general breeding technique and a general genetic technique, which are well-known to a person with ordinary skills in the art pertaining to the present invention, based upon information of amino acid sequences to be obtained by combining PPR motif and mitochondrion transfer sequence obtained based upon amino acid sequences described in SEQ ID NO 1 or SEQ ID NO 2 disclosed in this specification and amino acid sequences described in SEQ ID NO 3 or information of amino acid sequences described in SEQ ID NO 3.

Specifically by appropriate plant origin where the genes of the present invention are expressed, specifically by genus *Raphanus* plants including radish variety and related species thereof or other plants where the genome DNA including a cytoplasmic male sterile restoring gene has been transferred from these plants by techniques of hybridization or cell fusion; more specifically by (*Raphanus sativus* cv. Kosen), Ogura radish (*Raphanus sativus* L.) and Yuanhong radish (*Raphanus sativus* cv. Yuanhong) or genus *Raphanus* plants, such as these radish varieties and related species or techniques of hybridization or cell fusion, the genome DNA including a cytoplasmic male sterility restoring gene of these plants can be obtained, and for example, genes of the present invention can be acquired by isolating a DNA marker positioned in the vicinity of the Rf gene, by preparing a genome map showing the relationship of genetic distance between this DNA marker and the Rf gene, and by a positional cloning method (also referred to as chromosome walking) of the Rf region as a starting point of the genome map.

This method is started from discovering appropriate DNA marker on the genome DNA, and producing a genome map by measuring a genetic distance between the Rf gene and the DNA marker. It is necessary for the DNA marker to identify between the father-derived genome and the mother-derived genome, and it generally comprises the length of number 100 bp. Further, it is necessary for the DNA marker to place on the same chromosome, and a marker whose hereditary mode is substantially the same because the distance in between the gene is close, i.e., a marker that is genetically strongly linked is desirable.

The RFLP method has been conventionally used as the DNA marker isolation method; however, the RAPD method and AFLP (amplified fragment length polymorphism) method (Nucleic Acids Research, 1995, Vol. 23, No. 21, 4407-4414), which are simplified methods using PCR, are used. In particular, the AFLP method is effective as a means to obtain a marker that is genetically strongly linked. As a material to measure a genetic distance with a marker, an F<sub>2</sub> aggregate to be obtained by self-pollinating the F<sub>1</sub> generation where a recessive homo-individual not having a Rf<sub>1</sub> gene and a dominant homo-individual having the Rf<sub>1</sub> gene in the homo-individual and a BC<sub>1</sub> aggregate to be obtained by breeding the F<sub>1</sub> generation and the recessive homo-individual, which is a parent of the F<sub>1</sub> generation and does not have a target gene can be normally used.

As the recessive homo-individual, radish variety of cytoplasmic male sterile line, genus *Raphanus* plants including related species, to be more specific, kosena radish (*Raphanus sativus* cv. Kosena) and ogura radish (*Raphanus sativus* L.) in the cytoplasmic male sterile line; or genus *Brassica* plants where kosena radish (*Raphanus sativus* cv. Kosena)-derived cytoplasmic male sterility (kosena cms) or Ogura radish (*Raphanus sativus* L.)-derived cytoplasmic male sterility (ogura cms) is transferred, to be more specific, cms rapeseed can be used.

As the dominant homo-individual, radish variety, which is the Rf line, genus *Raphanus* plants including related species, to be more specific, kosena radish (*Raphanus sativus* cv. Kosena), Ogura radish (*Raphanus sativus* L.) and *Raphanus sativus*; or genus *Brassica* plants where genome DNA including a cytoplasmic male sterility restoring gene of genus *Raphanus* plants including these radish varieties and related species is transferred by a technique of hybridization or cell fusion, to be more specific, Rf rapeseed can be used.

For the F<sub>2</sub> aggregate to be obtained by self-pollinating the F<sub>1</sub> generation, which is obtained by breeding these parents, and the BC<sub>1</sub> aggregate to be obtained by breeding the F<sub>1</sub> generation and the recessive homo-individual, it is desirable to analyze 100 individuals normally and to analyze 1,000 individuals more preferably, and the more the number of individuals is increased, the greater the accuracy of the genome map becomes, and the physical distance from the DNA marker to the target gene becomes shorter. As similar to the case of the Rf gene, it becomes possible to obtain a DNA marker with shorter physical distance.

As a material for measuring the genetic distance between the DNA marker and the Rf gene, for example, thousands of F<sub>2</sub> aggregates obtained by self-pollinating the radish F<sub>1</sub> generation where cms line kosena radish (*Raphanus sativus* cv. Kosena) and Rf line Yuanhong radish (*Raphanus sativus* cv. Yuanhong) were bred based upon a method described in N. Koizuka, et al. Theor Appl Genet, 100:949-955, 2000 can be used. The DNA markers that are linked at positions with approximately 0.2 cM of genetic distance at both sides in the form of interposition of the Rf gene can be isolated by analyzing these, and a genome map showing the genetic distance between the marker and the Rf gene as shown in Fig. 1 can be prepared.

Following the preparation of the genome map, the genome DNAs corresponding to the positions are cloned, and it becomes possible to connect the DNA markers by sandwiching the target gene. Normally, since the

physical distance between the DNA marker and the target gene is great, the target gene region from the DNA marker shall be covered by connecting multiple clones having genome DNA fragments. The process to connect between these DNA markers with the clones is a contig production. Similar to the case of the Rf gene, the contig can be produced by connecting multiple clones having the genome DNA fragments so as to cover the Rf gene region in between the DNA markers existing close to the Rf gene.

An aggregate of the clones having the genome DNA fragments can be obtained by producing a genomic library. Normally, several types of vectors are used according to the length of clonable genome DNA, and for example, a library utilizing a lambda phage vector that can clone fragments up to approximately 20 kb, a cosmid vector that can clone comparatively longer fragments (up to 40 kb), and a BAC (bacterial artificial chromosome) that can clone longer fragments, such as, 100 kb or longer of fragments can be mentioned.

In any library, it is important that a value where the average length of the cloned fragments is multiplied by the number of aggregates in the library becomes a value, which is four or five times of the entire length (genome size) of genome supplied to the library. Since it is believed that the genome size of radish is approximately 500 Mbp, in the case that the average length is 20 kb as in the lambda phage vector, the number of aggregates becomes  $1.0 \times 10^5$  to  $1.25 \times 10^5$ , and in the case that the average length is 40 kb in the cosmid library, the number of aggregates becomes  $5.0 \times 10^4$  to  $6.25 \times 10^4$ . Since it is believed that the genome size of rapeseed is approximately 1,000 Mbp, in the case that the average length is 20 kb in the lambda phage vector, the number of aggregates becomes  $2.0 \times 10^5$  to  $2.5 \times 10^5$ , and in the case that the average length is 40 kb in the cosmid library, the number of aggregates becomes  $1.0 \times 10^5$  to  $1.25 \times 10^5$ .

The genome DNA to be provided to the library may be extracted from a living organism containing the target gene using a conventional method. In

the case of Rf gene, the radish variety, which is a Rf line, genus *Raphanus* plants including related species, to be more specific, kosena radish (*Raphanus sativus* cv. Kosena), ogura radish (*Raphanus sativus* L.) and Yuanhong radish (*Raphanus sativus* cv. Yuanhong), which are cytoplasmic male sterility restoring lines; or genus *Brassica* plants where the genome DNA containing a cytoplasmic male sterility restoring gene of genus *Raphanus* plants including these radish species and related species is transferred by a technique of hybridization or cell fusion, and to be more specific Rf rapeseed can be utilized. In general, it is believed that it is the most desirable that the genome DNA is extracted from a plant, which is the same Rf line as the parent that was utilized at the time of producing the F<sub>2</sub> aggregate or the BC<sub>1</sub> aggregate, and a genomic library is produced. The genome DNA can be prepared by following a conventional method, such as CTAB method (Murray, M. G. and Thompson, W. F. (1980) *NucleicAcids Res.*, 8, 4321).

In the contig production, at first, a clone maintaining DNA markers located at both sides of the Rf gene is isolated. The clone is isolated from the genomic according to a conventional method, using a plaque hybridization method in the case of the lambda phage library or using a colony hybridization method in the case of the cosmid library and the BAC library. Next, a contig is produced by isolating clones adjacent to the clone using the terminal region of the isolated clone as an index. After the production, the base sequence of the contig is determined according to a conventional method.

Due to the recent progress of the genome project, a technology to estimate functioning genes from base sequences of the genome DNA has been developing. The gene discovery program represented by "Genscan" can estimate genes with considerably high accuracy. Further, the homology search program represented by "BLAST" can estimate similarity of other genes and proteins. Utilizing this analysis software, target genes are estimated and isolated. Even in the case of the Rf gene, it is possible to

similarly isolate and identify genome DNA sequences of contig by utilizing similar analysis software. Further, analyzing them, a promoter part on the genome DNA base sequence, a structural gene part including intron, and a terminator part are expressed. Further, simultaneously, genes to be translated into a protein except for intron and amino acid sequences to the gene are expressed. As described above, it is possible to estimate the Rf genes on the contig with considerably high accuracy.

Further, in order to confirm whether or not what form the target genome is actually expressed within a living organism based upon the estimated gene sequence by utilizing the analysis software, mRNA is purified and it can be proved by isolating a supplementary DNA (cDNA) to this mRNA. Further, regarding from where the transcription has started, it is possible to prove by analyzing with a method where PCR is applied referred to as 5'-RACE method as a simplified method, and to be more certain, with the primer extension method or the S1 mapping method.

The above-mentioned methods are described in *Molecular Cloning: A laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989.

The isolated gene of the present invention based upon the base sequences estimated by the techniques specifically includes DNA shown in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18, and it also becomes possible to easily isolate cDNA from other plant origins by a general genetic technical based upon the DNA sequence.

Specifically by appropriate plant origin where the genes of the present invention are expressed, specifically by Genus *Raphanus* plants including radish variety and related species thereof or other plants where the genome DNA including a cytoplasmic male sterile restoring gene has been transferred from these plants by techniques of hybridization or cell fusion; more specifically by (*Raphanus sativus* cv. Kosena), Ogura radish (*Raphanus sativus* L.) and Yuanhong radish (*Raphanus sativus* cv. Yuanhong) or Genus

Raphanus plants, such as these radish varieties and related species or techniques of hybridization or cell fusion, a cDNA library is prepared in accordance with a usual method from the genus Brassica plants where the genome DNA including a cytoplasmic male sterility restoring gene of these plants, and cDNA equivalent to the gene of the present invention can be isolated using appropriate DNA fragments, which are unique to the gene of the present invention, as a probe, or by selecting the desired clones using an antibody to the translation product of the gene of the present invention.

In the above-mentioned description, as an origin of cDNA, various cells, tissues and derived cultured cells where the gene of the present invention is expressed are illustrated. Further, separation of entire RNA from these, separation and purification of mRNA, acquisition of cDNA and cloning can be all conducted in accordance with a conventional method.

The method of screening the gene of the present invention from the cDNA library is also not limited, but a normal method can be followed.

As the probe used herein, DNA, which is chemically synthesized based upon the information regarding the base sequences of the gene of the present invention, can be generally used; however, the gene of the present invention, which has already been acquired, and fragments thereof can be excellently used. Further, a sense primer and an anti-sense primer, which are set based upon the base sequence information of the gene of the present invention, can also be used as a probe for screening.

The nucleotide sequences of the sense primer and anti-sense primer used for the primer include partial nucleotide sequences corresponding to the DNA coding for the amino acid sequences shown in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19, and the one having at least 15 to 50 successive bases, preferably 20 to 30 successive bases can be mentioned. Alternatively, a positive clone itself having the sequences can be used as a probe.

On the occasion of acquiring the gene of the present invention, a technique, which is normally used for the isolation of genes, may be used by

combining with a DNA/RNA amplification method according to the PCR method or the RACE method represented by the 5'-RACE method.

The primer used on the occasion of adopting the PCR method can be appropriately set based upon the sequence information of the gene of the present invention, which has been clarified by the present invention, and this can be synthesized in accordance with a conventional method. Furthermore, for isolation and purification of amplified DNA/RNA fragments, a conventional method can be used as described above, and for example, it can be conducted with gel electrophoresis.

The gene of the present invention obtained as described above or various DNA fragments can determine their base sequences in accordance with a conventional method.

According to the gene of the present invention obtained as described above, a presence of the gene of the present invention in an individual or various tissues and whether or not the gene is expressed can be characteristically detected by utilizing partial or entire base sequences of the gene.

As described above, the gene of the present invention can include the DNA coding for the amino acid sequences shown, for example, in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19; however, the gene is not particularly limited to this, but it is needless to say, it also includes a homologue of the gene.

Herein, the homologue of a gene means a series of related genes, which has a sequence homogeneity with the gene of the present invention (or its gene product), and which is recognized as one gene family according to the similarity of the structural characteristic and the biological function as described above, and it is needless to say, it includes an allelic gene of the gene, as well.

For example, the gene of the present invention is not limited to the gene having specific base sequences shown with SEQ ID NO 1 or SEQ ID NO

2, but it is also possible to have the base sequences to be selected by combining any codon to each amino acid residue shown with SEQ ID NO 3. Similarly, the gene of the present invention is not limited to the gene having specific base sequences shown with SEQ ID NO 15 or SEQ ID NO 16, but it is also possible to have the base sequences to be selected by combining any codon to each amino acid residue shown with SEQ ID NO 17, and the gene of the present invention is not limited to the gene having specific base sequences shown with SEQ ID NO 18, but it is also possible to have the base sequences to be selected by combining any codon to each amino acid residue shown with SEQ ID NO 19. The selection of codon can be in accordance with a conventional method, and for example, frequency in use of codon of a host to be utilized can be taken into consideration.

Further, as described above, the gene of the present invention even includes DNA hybridizing a DNA having base sequences shown in SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 15, SEQ ID NO 16, SEQ ID NO 18 or a part of them. This DNA is a DNA having a certain level or more of homology with the DNA having base sequences shown in SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 15, SEQ ID NO 16, SEQ ID NO 18 or a part of them.

The above-mentioned DNA having a certain level or more of homology is polynucleotide or a complementary strand polynucleotide having at least 70 % of identity, preferably at least 90 % of identity, or more preferably at least 95 % of or further more preferably 97 % of identity with base sequences shown in SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 15, SEQ ID NO 16 or SEQ ID NO 18, or a part of them, amino acid sequences shown in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19, or base sequences coding a part of them.

To be more specific, for example, the DNA having base sequences shown in SEQ ID NO 1, SEQ ID NO 2, SEQ ID NO 15, SEQ ID NO 16 or SEQ ID NO 18, or a part of the base sequences and the DNA having base sequences to be hybridized can be illustrated under stringent conditions with

0.2 × SSC at 60 degrees C containing 0.1 % SDS or 1 × SSC at 60 degrees C containing 0.1 % SDS.

【 0 0 6 8 】

Further, among the DNAs of the present invention, in particular, DNA having base sequences where one to multiple bases are deleted, added and/or substituted in base sequences or a part of them described in SEQ ID NO 1 or SEQ ID NO 15, and participating in the restoration from sterility of cytoplasmic male sterile individual to fertility;

DNA having base sequences where one to multiple bases are deleted, added and/or substituted in base sequences or a part of them described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18, and participating in the restoration from sterility of cytoplasmic male sterile individual to fertility; and

DNA having base sequences where one to multiple bases are deleted, added and/or substituted in base sequences or a part of them described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19, and participating in the restoration from sterility of cytoplasmic male sterile individual to fertility; can be produced by an optional method, which is well-known to a person with ordinary skills in the art pertaining to the present invention, such as chemical synthesis, genetic technique or induction of mutation. For example, a mutant gene can be acquired by inducing mutation to DNAs by utilizing the DNA having the base sequences or a part of the base sequences described in SEQ ID NO 1, 2, 15, 16 or 18.

As a method to obtain the mutant gene, for example, well-known methods, such as a method using a random mutant, targeted mutant or synthetic gene (see New Genetic Handbook, Supplementary Edition of Experimental Medicine, Yodosha Co., Ltd., 1996) can be used.

Specifically, a method to bring into contact and action with a drug to be mutagens, a method to radiate ultraviolet rays, or a genetic technique can be used upon the DNA having the base sequences described in SEQ ID NO 1

or 2 or part of the base sequences. Since the site-directed mutagenesis, which is one of genetic techniques, is a technique where a specific variation can be introduced to a specific position, it is useful, and it can be conducted in accordance with the method described in Molecular Cloning Second Edition.

(4) Vector containing DNA of the present invention:

The DNA of the present invention can be used as a recombinant vector by recombining into an appropriate vector. The type of vector may be an expression vector or non-expression vector, and it can be selected based upon the purpose.

As a cloning vector, a vector that can be autonomous-replicated among *Escherichia coli* K12 strains is preferable, and a vector for *Escherichia coli* expression, such as a phage vector or plasmid vector that can be used, may be used as a cloning vector. Specifically, it can include ZAP Express [Stratagene, Strategies, 5, 58 (1992)], pBluescript II SK(+) [Nucleic Acids Research, 17, 9494 (1989)], Lambda ZAP II ((manufactured by Stratagene),  $\lambda$ gt10,  $\lambda$ gt11 [DNA Cloning, A Practical APPROach, 1, 49 (1985)],  $\lambda$ TriplEx (manufactured by CloneTech),  $\lambda$ ExCell (manufactured by Pharmacia), pT7T318U (manufactured by Pharmacia), pcD2 [Mol. Gen. Bio1., 3, 280 (1983)], pMW218 (manufactured by Waco Pure Chemical Industries, Ltd.), pUC118 (manufactured by Takara Shuzo Co., Ltd.), pEG400 [Bac., 172, 2392 (1990)], and pQE-30 (manufactured by QIAGEN).

The expression vector can be selected by considering a combination with a host, and preferably, a vector that enables autonomous replication or incorporation into a chromosome, and that contains a promoter at the position where the gene of the present invention can be transcribed can be used.

When a bacillus is used as a host cell, it is preferable that the expression vector for expression of a DNA is autonomously-replicable in the bacillus; simultaneously, it is preferable that the vector is a recombinant vector comprising a promoter, a liposome-binding sequence, the DNA and the

transcription termination sequence. A gene that controls the promoter may be contained.

The expression vector for bacillus, for example, pBTrp2, pBTac1, pBTac2 (all are commercially supplied by Boehringer Mannheim GmbH), pKK233-2 (manufactured by Pharmacia), pSE280 (manufactured by Invitrogen), pGEMEX-1 (manufactured by Promega), pQE-8 (manufactured by QIAGEN), pQE-30 (manufactured by QIAGEN), pKYP10 (Japanese Unexamined Patent Application S58-110600), pKYP200 [Agric. Biol. Chem., 48, 669 (1984)], PLSA1 [Agric. Biol. Chem., 53, 277 (1989)], pGEL1 [Proc. Natl. Acad. Sci. USA, 82, 4306 (1985)], pBluescriptII SK+, pBluescriptII SK (-) (manufactured by Stratagene), pTrs30 (FERMBP-5407), pTrs32 (FERM BP-5408), pGEX (manufactured by Pharmacia), pET-3 (manufactured by Novagen), pTerm2 (US4686191, US4939094, US5160735), pSupex, pUB110, pTP5, pC194, pUC18 [Gene, 33, 103 (1985)], pUC19 [Gene, 33, 103 (1985)], pSTV28 (manufactured by Takara Shuzo Co., Ltd.), pSTV29 (manufactured by Takara Shuzo Co., Ltd.), pUC118 (manufactured by Takara Shuzo Co., Ltd.), pPA1 (Japanese Unexamined Patent Application S63-233798), pEG400 [J. Bacteriol., 172, 2392 (1990)], pQE-30 (manufactured by QIAGEN) can be illustrated. The promoter for bacillus includes, for example, promoters derived from *Escherichia coli* or phage, such as a trp promoter (P trp), a lac promoter (P lac), a PL promoter, a PR promoter or a PSE promoter; a SP01 promoter, a SP02 promoter and a penP promoter.

As the expression vector for fermentum, for example, a promoter, such as YEp13 (ATCC37115), YEp24 (ATCC37051), Ycp50 (ATCC37419), pHS19 and pHS15 are exemplified. A promoter for fermentum includes, for example, a PHO5 promoter, a PGK promoter, a GAP promoter, an ADH promoter, a gal1 promoter, a gal10 promoter, a heat shock protein promoter, an MFa1 promoter or a CUP1 promoter.

As the expression vector for animal cells, for example, pcDNA1, pcDM8 (commercially supplied from Funakoshi Corporation), pAGE107 [Japanese

Unexamined Patent Application H3-22979; Cytotechnology, 3, 133, (1990)], pAS3-3 (Japanese Unexamined Patent Application H2-227075), pCDM8 [Nature, 329, 840, (1987)], pcDNAI/AmP (manufactured by Invitrogen), pREP4 (manufactured by Invitrogen), pAGE103 [J.Biochem., 101, 1307 (1987)], and pAGE210 can be illustrated. The promoter for animal cells includes, for example, a promoter of IE (immediate early) gene of cytomegalovirus (human CMB), an initial promoter of SV40, a retrovirus promoter, a metallothionein promoter, a heat shock promoter and a SRA promoter.

As the expression vector for plant cells, for example, pIG121-Hm [Plant Cell Report, 15, 809-814 (1995)], pBI121 [EMBO J. 6, 3901-3907 (1987)], pLAN411 and pLAN421 (Plant Cell Reports 10 (1991), 286-290) can be illustrated. Further, when, in particular, 10 kb or longer DNA fragments are induced into plants, it is desirable to use improved vectors so as to stably maintain and introduce long-chain DNA. For example, pBIBAC2 (Gene 200 (1997) 107-116), pYLAC7 (PNAS 96 (1999) 6535-6540) and pBGRZ2 (Bioscience and Industry, 55 (1997), 37-39) are mentioned.

The promoter for plant cells includes, for example, a cauliflower mosaic virus 35S promoter [Mol.Gen.Genet (1990) 220, 389-392]. Furthermore, details of plant transformation will be separately described later.

(5) Transformant having DNA of the present invention:

The transformant having DNA of the present invention can be produced by introducing the above-mentioned recombinant vector (preferably, expression vector) to a host.

Specific examples of host cells of bacillus includes microorganisms belonging to Escherichia, Corynebacterium, Brevibacterium, Bacillus, Microbacterium, Serratia, Pseudomonas, Agrobacterium, Alicyclobacillus, Anabaena, Anacystis, Arthrobacter, Azobacter, Chromatium, Erwinia, Methylobacterium, Phormidium, Rhodobacter, Rhodopseudomonas,

Rhodospirillum, Scenedesmus, Streptomyces, Synnecoccus or Zymomonas. A method to introduce a recombinant vector into a cell host includes, for example, a method using calcium ions or the protoplast method.

Specific examples of fermentum host can include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces lactis, Trichosporon pullulans and Schwanniomyces alluvius.

As a method for introducing a recombinant vector into fermentum host, any methods for inducing DNA into fermentum can be used, and for example, they can include an electroporation method, a spheroplast method and a lithium acetate method.

The animal cell host can include Namalwa cells, COS1 cells, COS7 cells and CHO cells.

As an introduction method of a recombinant vector into animal cells, any method for introducing DNA into animal cells can be used, and for example, the electroporation method, a calcium phosphate method and a lipofection method can be used.

A transformant using plant cells will be described later.

(6) Acquisition method for proteins of the present invention:

The acquisition method for the proteins of the present invention is not particularly limited. The proteins of the present invention can be isolated, expressed or synthesized by utilizing a general genetic technique for a person with ordinary skills in the art pertaining to the present invention, based upon information about amino acid sequences obtained by combining the information about the amino acid sequences described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19 or PPR motifs or mitochondrion transfer sequences obtained based upon the amino acid sequences described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19.

For example, the expression can be realized by isolating or synthesizing DNA that codes for the protein of the present invention and by introducing the DNA into cells.

The proteins of the present invention can be obtained, for example, by culturing transformant having the genes of the present invention, by producing and accumulating the proteins of the present invention in the culture, and by collecting the proteins from the culture.

The method for culturing the transformant having the genes of the present invention can be in accordance with the normal method used for the culture of hosts.

When the transformant of the present invention is prokaryote, such as *Escherichia coli*, or eukaryote, such as yeast fungus, the culture medium for culturing these microorganisms may be either natural or synthetic culture medium, as long as it is the culture medium that contains a carbon source, a nitrogen source or inorganic salts where the microorganisms can be utilized, and where the culture of the transformant can be efficiently conducted. It is preferable to culture the host under aerobic conditions, such as shaking culture or deep ventilation spinner culture, and the culture temperature is 15 degrees C to 40 degrees C, and the culture time is normally for 16 hours to 7 days. During culturing, pH is maintained at 3.0 to 9.0. The pH is adjusted by inorganic or organic acid, an alkali solution, urea, calcium carbonate or ammonia. Further, antibiotics, such as ampicillin or tetracycline, may be added during the culturing as occasions demand.

As the culture medium where the transformant obtained using animal cells as a host cell, generally-used RPM11640 culture medium [The Journal of the American Medical Association, 199, 519 (1967)], Eagle MEM culture medium [Science, 122, 501 (1952)], DMEM culture medium [Virology, 8, 396 (1959)], 199 culture medium [Proceeding of the Society for the Biological Medicine, 73, 1 (1950)] or culture media where cow's fetus serum is added into these culture media are used. The culture is normally conducted for 1

day to 7 days under conditions of 6 to 8 of pH, 30 degrees C to 40 degrees C of temperature and in the presence of 5 % CO<sub>2</sub>. Further, antibiotics, such as Kanamycin or penicillin, may be added to the culture medium during the culturing as occasions demand.

As the culture medium for culturing the transformant obtained using plant cells as a host cell, culture media normally used according to plant species, such as MS culture medium or R2P culture medium. The culturing is normally conducted for 1 day to 21 days under conditions at 6 to 8 of pH and 15 degrees C to 35 degrees C [of temperature]. Further, antibiotics, such as kanamycin or hygromycin, may be added to the culture medium during the culturing as occasions demand.

In order to isolate and purify the proteins of the present invention participating in the restoration from sterility of cytoplasmic male sterile individual to fertility from the culture of the transformant, normal protein isolation and purification methods may be used.

For example, when the proteins of the present invention are expressed in the state where they are dissolved into cells, after the culturing, the cells are collected via centrifugal separation and suspended into an aquatic buffer solution, the cells are crushed by an ultrasonic crusher, a French press, a Menton-Gaulin homogenizer or a dyno-mill, and a cell-free extract is obtained. A purified preparation can be obtained from supernatant obtained via the centrifugal separation of the cell-free extract using techniques, such as a normal protein isolation purification method, i.e., a solvent extraction method, a salting-out method with ammonia sulfate, a desalination process, a precipitation method with an organic solvent, an anion exchange chromatography method using a resin, such as diethylaminoethyl (DEAE) sepharose, DIAION HPA-75 (manufactured by Mitsubishi Chemical), a cation exchange chromatography method using a resin, such as S-Sepharose FF (manufactured by Pharmacia), a hydrophobic chromatography using a resin, such as butyl sepharose or phenyl sepharose, a gel filtration method using

molecular sieving, an affinity chromatography method, a chromatography focusing method, singularly or in combination.

Further, when the proteins form an insolubilized body within cells and it is expressed, similarly the cells are crushed after they are collected, and after the proteins are collected using a normal method from a precipitation fraction obtained via the centrifugal separation, the insolubilized body of the proteins is solubilized with a protein denaturant. After the solubilized solution is diluted or dialyzed to a solution, which does not contain a protein denaturant or a dilute solution where the concentration of the protein denaturant is low enough not to cause denaturing of the proteins, and after the proteins are constructed to a normal steric structure, a purified preparation can be obtained using the similar isolation and purification method.

When the proteins of the present invention or derivatives thereof, such as sugar modified bodies, are secreted to the outside of the cell walls, the proteins or derivatives thereof, such as glycosylation bodies, can be collected from the culture supernatant. In other words, a soluble fraction is acquired by processing the culture using the similar technique, such as centrifugal separation, and a purified preparation can be obtained from the soluble fraction by using the similar isolation and purification method.

Further, the proteins of the present invention can be produced using a chemical synthesis method, such as an Fmoc method (fluorenylmethyloxycarbonyl method) or a tBoc method (t-butyloxycarbonyl method). Further, the proteins of the present invention can also be synthesized by utilizing peptide synthesizers by Sowa Trading (manufactured by US Advanced Chem Tech), Perkin Elmer Japan (manufactured by US Perkin Elmer), Amersham Pharmacia Biotech (manufactured by Amersham Pharmacia Biotech), Aloka (manufactured by US Protein Technology Instrument), Kurabo (manufactured by US Synthecell-Vega), Japan PerSeptive Limited (manufactured US PerSeptive), Shimadzu Corporation.

(7) Transformant of plants having DNA of the present invention:

The base sequences described in SEQ ID NO 1 and SEQ ID NO 15 are base sequences in the extracted form of plant genome-originated base sequences. These base sequences contain a promoter and terminator required for the gene expression to be operable. The vector to be introduced can clone the genes to a general cloning vector, such as cosmid pWE15 (manufactured by STRATAGENE), in the case of a direct introduction method. In the case of utilizing *Agrobacterium*, it can be cloned to a general vector for plant transformation, such as pBI121 (manufactured by Clontec).

Further, DNA of the base sequences where a part of intron is extracted from this sequence (genomic sequence), DNA of base sequences where almost all of intron is extracted, DNA shown with SEQ ID NO 2 or portions corresponding to the 238th to 2,064th base sequences, DNA shown with SEQ ID NO 16 or portions corresponding to the 238th to 2,064th base sequences, DNA shown with SEQ ID NO 18 or portions corresponding to the 244th to 2,073rd base sequences; or DNA that codes for proteins shown with SEQ ID NO 3 or portions corresponding to 80 to 687 residues, with SEQ ID NO 17 or portions corresponding to 80 to 687 residues, or with SEQ ID NO 19 or portions corresponding to 82 to 690 residues may be introduced into plant cells.

Herein, DNA having the 3,754th to 5,091st base sequences among the base sequences described in SEQ ID NO 1 or having the 1st to 811th base sequences among the base sequences described in SEQ ID NO 15 is promoter having a capability to transfer mRNA with anther, and it is preferably used for restoring male sterility.

In addition, promoter and terminator portions may be substituted by a promoter and a terminator that function in the well-known plant cells.

Furthermore, when DNA shown with SEQ ID NO 2 or with the 238th to 2,064th DNA thereof, DNA shown with SEQ ID NO 16 or with the 238th to

2,064th DNA thereof and DNA shown with SEQ ID NO 18 or with the 80th to 687th DNA thereof or DNA that codes for the proteins shown described in SEQ ID NO 3 or with portions corresponding to the 80 to 687 residues, SEQ ID NO 17 or with portions corresponding to the 80 to 687 residues or SEQ ID NO 19 or with portions corresponding to the 82 to 690 residues is introduced into the plant cells, a promoter and a terminator are required other than this DNA. As the general vectors normally used well, pBI121 (manufactured by Clontec) is mentioned; however, in this vector, a 35S promoter of cauliflower mosaic virus is used for the promoter and a terminator of nopaline synthesis enzyme existing in Ti plasmid of *A. tumefaciens* is used for the terminator. Further, as the promoter required for the expression, the 35S promoter of cauliflower mosaic virus is not limited to be used but a *rbcs* promoter widely existing in plants may be used, and it is more preferable to use a type of promoter that is expressed during the growing period of pollen, such as T29 promoter, and it is further preferable to use the original promoter coordinated at the upstream of the gene. The terminator is also not limited to the terminator of the nopaline synthesis enzyme, but a 35S terminator of cauliflower mosaic virus can be used, and it is more preferable to use an original terminator coordinated at the downstream of the gene.

In addition, when the DNA described in SEQ ID NO 2 or with the 238th to 2,064th base [sequences] thereof, the DNA described in SEQ ID NO 16 or with the 238th to 2,064th base [sequences] thereof or the DNA described in SEQ ID NO 18 or with the 244th to 2,073rd base [sequences] thereof, or the DNA that codes for the proteins described in SEQ ID NO 3 or with the portions corresponding to 80 to 687 residues thereof, in SEQ ID NO 17 or with the portions corresponding to 80 to 687 residues thereof or in SEQ ID NO 19 or with the portions corresponding to 82 to 690 residues thereof are used for the purpose of restoring from the sterility of cytoplasmic male sterile individual to fertility, a transfer sequence to mitochondria becomes required, in addition to these sequences.

As the transfer sequence to mitochondria, DNA shown with the 1st to 237th base [sequences] in SEQ ID NO 2 or the 1st to 79th amino acid sequences in SEQ ID NO 3, DNA shown with the 1st to 237th base [sequences] in SEQ ID NO 16 or the 1st to 79th amino acid sequences in SEQ ID NO 17 or DNA shown with the 1st to 243RD base [sequences] in SEQ ID NO 18 or the 1st to 81st amino acid sequences in SEQ ID NO 19 or any other well-known transfer sequences can be utilized.

The inventors of the present application produced a vector for plant transformation for the purpose of introducing DNA of the Rf gene containing intron, which is contained from the original promoter to terminator existing in the genome, in the original form, shown in SEQ ID NO 1 in the embodiments described below. After the base sequences shown in SEQ ID NO 1 were cut off by a restricted enzyme from clones comprising a part of contig, they were sub-cloned to appropriate cloning vectors, and then, fragments sub-cloned to the vectors for plant transformation pKM424 and pBGRZ2 were introduced, and a vector that can introduce the fragments into plants was obtained. This vector was introduced to *Agrobacterium* for plant transformation. The DNA fragments are incorporated into the plant genome by infecting this *Agrobacterium* where the vector is maintained to plants.

For the plants where the gene of the present invention is applied, for example, oil crops, such as rapeseed, sunflower, soybean or palm; for example, cereal crops, such as rice, corn or wheat; for example, flowers and ornamental plants, such as tobacco or petunia; and for example, various vegetables, such as tomato, broccoli, cabbage, Chinese cabbage or carrot, can be exemplified.

Among them, genus Brassica plants, such as rapeseed, cabbage, Chinese cabbage or broccoli, and tomatoes are preferable, and rapeseed, cabbage, Chinese cabbage and broccoli are particularly preferable, and rapeseed is the most preferable.

In this specification, the transformation plant source includes seeds, germs, nursery plants, callus, cultured cells and plant bodies, and a

preferable part may be selected according to a subject plant as normally conducted by a person with ordinary skills in the art pertaining to the present invention, for example, in the case of rapeseed, germ or protoplast; in the case of soybeans, callus or cultured cells; in the case of sunflower, germ; in the case of palm, callus or cultured cells; in the case of rice, germ, callus, cultured cells or protoplast; in the case of corn, germ, nursery plants, callus, cultured cells or protoplast; in the case of wheat, germ, callus or cultured cells; in the case of cabbage or broccoli, germ, callus, cultured cells or protoplast; and in the case of carrot, germ, callus, cultured cells or protoplast.

The transformation procedure to plants can be conducted in accordance with a conventional method, and for example, a method where after a vector is once introduced into *Agrobacterium*, the vector is introduced by infecting the *Agrobacterium* with plant cells, or another method for directly introducing the vector into cells using an electroporation method, a DEAE dextran method, a calcium phosphate method, a polyethylene glycol method, or a particle gun method can be exemplified.

For example, as a preferable gene introduction method in the case of rapeseed, the following methods are exemplified:

Hypocotyl of rapeseed variety sprouted in the MS culture medium containing sugar, such as sucrose, as a carbon source is pre-cultured on the MS culture medium containing 2,4-dichlorophenoxyacetic acid and sucrose. *Agrobacterium* proliferated in the YEB culture medium is harvested via centrifuge, and re-suspension is conducted in the MS culture medium containing sucrose. After the above-mentioned rapeseed hypocotyl is added to this suspension and the mixture is concussed, the extracted hypocotyl is put back into the original pre-cultured medium and it is co-cultivated for three days, and then, the culture is transferred to a selection medium containing plant hormone, such as zeatin or benzylaminopurine, carbenicillin and kanamycin, and selection is conducted. A reproduced individual can be obtained by culturing the obtained regeneration bud via this process in an

elongation culture medium optionally containing plant hormone, such as benzylaminopurine, and consecutively in a rooting culture medium optionally containing plant hormone, such as naphthalene acetic acid or benzylaminopurine), and F1 variety where the fertility is restored can be obtained by breeding this individual with a cms line individual.

As described above, it becomes possible to restore from the sterility of the cytoplasmic male sterile individual to the fertility by introducing the DNA of the present invention into plants. Furthermore, for the regenerated individual, the expression can be confirmed by breeding with cms line rapeseed and by investigating the fertility of descendant thereof; however, in the case of transformation using rapeseed having a cms cytoplasm as a material, the pollen fertility can be investigated by transferring a transformant (regenerated individual) where a root is formed as described above into soils containing normal fertilizers and by blooming, and this is preferable because the method is temporally and operationally simple.

Further, in the transformation, in the case of transformation as described above by utilizing cells or tissues of rapeseed having the cms cytoplasm, preferably, hypocotyl, cotyledon, leaves, pollen, cultured cells, callus or protoplast, as cells to be used: a plant individual where the pollen fertility is restored can be obtained by transferring the plant body (regenerated individual) obtained with the above-mentioned method to soils containing normal fertilizers and by blooming.

In other words, using cms cells, after the DNA of the present invention is introduced using the above-mentioned gene introduction method to the cms cells and the cells where the DNA is incorporated into a nuclear are selected using resistance to antibiotics or herbicide-tolerant selection markers, such as kanamycin, as an index, a plant body where the DNA is incorporated into the nuclear can be obtained by culturing in the prolongation culture medium or the rooting culture medium. The male sterile character is restored, and this plant body becomes fertile.

As a method for detecting the gene participating in the restoration from cytoplasmic male sterility, it can be conducted by confirming such that a quantity of base sequences to be amplified by the primer or a quantity of base sequences to be detected by the probe is one gene or more per genome in the target biological sample using an oligonucleotide primer with 15 mer to 50 mer optionally set from the DNA according to any of Claims 1 to 4 or a probe with at least 15 mer or more comprising entire or a part of the DNA according to any of Claims 1 to 4.

The specific confirmation technique includes, for example, the PCR method and the Southern hybridization method, and between them, the PCR method is preferable. These techniques can be conducted in accordance with the method described in *Molecular Cloning: A laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989 (hereafter, abbreviated as "Molecular Cloning 2nd Ed").

In order to confirm that one gene or more exist in one genome, it is necessary to confirm the same degree of amplification using the same number of DNA copies as a template, as a simplified method in the PCR method, and to be more accurate, using the quantification PCR method, this can be confirmed by comparing the amplification quantity of known genes with a quantity of base sequences to be amplified by the primer in the biological sample, using any primer that amplifies known gene, which is known that one gene exists in one genome as internal standard. In the Southern hybridization method, DNA of the fertility restored line plant individual, which is known that one gene of DNA exists in one genome, is compared to the DNA of a target plant sample, and whether or not the detected DNA quantity is the same or greater.

As the primer used for the PCR method, for example, oligonucleotide with 15 mer to 50 mer, which is the same as DNA sequences described in SEQ ID NO 1 or SEQ ID NO 2, or which has complementation, is exemplified.

As the probe used for the Southern hybridization method, the entire area of double-strand DNA or a portion with at least 15 mer or more of DNA sequences described in SEQ ID NO 1 or SEQ ID NO 2, or single-strand DNA or the entire area of complementation or a portion with at least 15 mer or more are exemplified. Further, as described above, DNA having a certain level or more of homology with the base sequences of DNA used as the probe as mentioned above can be exemplified. The 'a certain level or more of homology' herein means, for example, 70 % or more, preferably 80 % or more, more preferably 90 % or more, further preferably 93 % or more, particularly preferably 95 % or more, and the most preferably 97 % or more. Furthermore, as DNA having a certain level or more of homology herein contains both polynucleotide having the homology and complementary strand polynucleotide.

The method for detecting genes as described above is not only used for confirmation about whether or not the DNA is incorporated in the transformant, but also used as a means to confirm the presence of Rf gene in the individual where the introduction of Rf gene is attempted due to breeding. If using this method, when the Rf gene is introduced into a cytoplasmic male sterile individual, it is possible to confirm the presence of the Rf gene before blooming. Further, when the Rf gene is introduced into an individual having normal cytoplasm, the fertility of the next generation individual, which was obtained by breeding pollen at the time of blooming to the cytoplasmic male sterile individual, has to be confirmed; however, the presence of the Rf gene can be confirmed before this stage by using this method. This utilization method is generally referred to as utilization of marker DNA or marker DNA breeding. In the case of Rf gene, utilization as a marker DNA (Rf marker) of the Rf gene can be considered. The Rf marker, as described above, is important using a plant where the Rf gene is introduced by breeding, which is the recombinant where the DNA is introduced or non-recombinant, and breeding the practical variety.

Furthermore, in order to confirm whether or not the introduced DNA functions as the Rf gene, it is also possible to confirm the fertility restoration of the transformant; however, this can be conducted using the method mentioned below.

As described above, the Rf gene restores the fertility of a plant by reducing the accumulation within the mitochondria of ORF125 or ORF138, which is a cms causal protein. Therefore, in the mitochondria of the transformant, whether or not the introduced gene is the Rf gene can be confirmed by confirming the reduction in the accumulation of ORF125 or ORF138.

The method to confirm the reduction in accumulation of ORF125 or ORF138 protein in the mitochondria is a method to confirm that when ORF125 or ORF138 protein is detected using the Western blotting method in accordance with the conditions described in this specification, a signal amount of an antibody to the mitochondria genome-derived protein to be used as an internal standard, for example, anti-F<sub>1</sub>-F<sub>0</sub>ATPase (hereafter, abbreviated as ATPA) described in N. Koizuka, et al. Theor Appl Genet, 100: 949-955, 2000 is equal to the cytoplasmic male sterile individual and the transformed individual where the DNA is introduced to a fertility restoration individual or cytoplasmic male sterile individual, and, the accumulation in the transformed individual where the DNA is introduced into the fertility restored individual or the cytoplasmic male sterile individual is reduced by 50 % or more, preferably 60 % or more, and further preferably 80 % or more, compared to the accumulation of ORF125 or ORF138 protein in the cytoplasmic male sterile individual.

In actuality, it is observed that in buds of cytoplasmic male sterile radish having ORF125, when the fertility restoring gene Rf is introduced, the accumulation of ORF125 protein is drastically decreased, and it is hardly detected. Further, regarding rapeseed, in the buds of the fertility restored rapeseed where the fertility restoring gene is introduced into the cytoplasmic

male sterile rapeseed having cytoplasm ORF125 due to breeding; it is observed that the accumulation of ORF125 protein is reduced by 80 % or more. Further, even in embryos, in the buds of transformed rapeseed where the DNA is introduced to the cytoplasmic male sterile individual, it is also observed that the accumulation of ORF125 protein is reduced by 80 % or more.

Furthermore, the antibody to ORF125 or ORF138 protein in the method can be obtained by using the general technique as described below. In other words, antiserum can be obtained by immunizing these proteins to animals as an antibody, and an immune globulin G antibody can be purified by using an affinity column where protein A is combined. An antigen to be used can be obtained by purifying proteins from expressed cytoplasmic male sterile plants or cultured cells thereof using a usual method. Further, this antigen can be obtained by connecting ORF125 or ORF138 gene to the expression vector for expressing in *Escherichia coli* or ferment and by similarly purifying. In addition, peptide where the entire length or a part of ORF125 or ORF138 is chemically synthesized can be used as an antigen. The antibody to ATPA can also be obtained using a similar technique.

In addition, a new hybrid seed production system that does not require any male sterility maintainer line (maintainer line) [*sic.*], which is necessary for the hybrid seed production, can be produced by specifically and temporarily controlling the expression of the DNA of the present invention, by introducing a part of or all of genes of the present invention along with an induced promoter into cells having cms cytoplasm and having the DNA of the present invention.

In other words, the cms line rapeseed is normally sterile, in order to proliferate and maintain the cms line, a maintainer line where cms and Rf do not participate is separately required, and conventionally, three lines of plants, Rf line, cms line and maintainer line, are required for the purpose of production of hybrid seeds; however, because the Rf gene is isolated and

identified according to the present invention, on the occasion of the hybrid production, the cms line that enables proliferation and maintenance without the maintainer line can be structured by inducing a promoter by a chemical substance and using a method for controlling the expression of restoring genes.

Specifically, a part of or an entire length of the gene of the present invention is incorporated into a promoter induced from the outside, for example, a vector having a promoter with drug sensitivity in an antisense or sense direction, and cells having the cms cytoplasm and having the DNA of the present invention are transformed using the vector.

As the cells having the cms cytoplasm and having the DNA of the present invention, not only cells having the cms cytoplasm are transformed by the DNA of the present invention in accordance with the above-mentioned method, but cells obtained by breeding the cms line and the Rf line are also usable.

As the inducible promoter, for example, it is known in Japanese Unexamined Patent Application H6-46697, and as a method for vector production and transformation, the similar techniques can be exemplified.

Since the transformant, which is a cell having the cms cytoplasm obtained by the above-mentioned method and having the DNA of the present invention, and where a part or entire DNA of the present invention is incorporated along with the inducible promoter, normally does not induce a promoter, the plant shows fertility because of the originally existing Rf genes and the line can be maintained by self-pollinating; however, on the occasion of hybrid production, if a chemical substance having a capability to induce a promoter acts on this plant and the promoter is induced, the expression of the Rf gene is inhibited. Because of this action, the plant becomes male sterile, and this can be used as the cms line at the time of hybrid seed production.

Therefore, if this method is used, since proliferation and maintenance can be conducted by self-pollinating even with the cms line, the maintainer

line is no longer required even though three lines used to be conventionally required for the production of hybrid seeds; therefore, it becomes possible to drastically reduce the production cost.

Furthermore, the details disclosed in each specification of Japanese Patent Application 2001-12800, Japanese Patent Application 2001-202082 and Japanese Patent Application 2001-20083, which are basis of the claim of priority of the present application, shall be all disclosed in the present specification by citation.

Hereafter, embodiments will be described in further detail; however, the present invention shall not be limited to these embodiments.

#### Embodiment

Embodiment 1: Isolation of DNA marker linked to cytoplasmic male sterility restoring gene and preparation of genome map

In order to an isolate fertility restoring gene (Rf gene), first, it is necessary to isolate the DNA marker positioned in the vicinity of the Rf gene and to prepare a genome map showing the gene distance relationship between this DNA marker and the Rf gene. As a starting point, positional cloning of the Rf region was conducted.

For the DNA marker isolation method, AFLP was conducted in accordance with an AFLP analysis system I AFLP starter primer kit, based upon the AFLP (amplified fragment length polymorphism) method (Nucleic Acids Research, 1995, Vol.23, No.21 4407-4414). As materials to measure the genetic distance from the marker, approximately 2,100 individuals of F2 aggregation obtained by self-pollinating eight individuals of radish F1 generation, which were [obtained] by breeding one individual of cms line kosena radish (*Raphanus sativus* cv. Kosena) ((KC2/KA1)-1) and one individual (yuan 10-3) of Rf line Yuanhong radish (*Raphanus sativus* cv. Yuanhong) in accordance with the method described in N. Koizuka, et al. Theor Appl Genet, 100:949-955 2000, were used. As a result, five markers

linked to the positions separated at 0.2 to 0.3 cM of genetic distance at both sides so as to interpose the Rf gene were isolated. A genome map showing the genetic distance between the DNA marker and the Rf gene is shown in Fig. 1, respectively.

Embodiment 2: Preparation of contig based upon genome map and analysis of Rf gene

Following the preparation of the genome map, it is necessary to clone the genome DNA corresponding to the position, and to connect the DNA markers interposing the Rf gene. Herein, since the distance between the DNA marker and the Rf gene is apart, the contig of the Rf gene region to cover in between the DNA markers was prepared by connecting multiple clones having the genome DNA fragments.

An aggregation of clones having genome DNA fragments is referred to as a genomic library, and we prepared two types of libraries. As a DNA donor, the genome DNA was prepared using the CTAB method (Murray, M. G. and Thompson, W. F. (1980) *Nucleic Acids Res.*, 8, 4321) with Yuanhong radish (*Raphanus sativus* cv. Yuanhong), which is the same as a parent of the restoration line utilized at the time of production of F<sub>2</sub> aggregation. For the library, a lambda phage library with 20 kb of average length and  $1.5 \times 10^5$  of the number of aggregations was prepared using a  $\lambda$ DASHII vector (manufactured by STRATAGENE) as a lambda vector. Further, a cosmid library with 40 kb of average length and  $5.5 \times 10^4$  of the number of aggregations was prepared using a pWEB::TNC vector (manufactured by EPICENTRE TECHNOLOGIES) as a cosmid vector.

For the contig production, first, lambda clones were isolated from the prepared lambda library using the DNA markers positioned at both sides of the Rf gene as indexes, using the plaque hybridization method. Further, the cosmid clones were isolated from the cosmid library using the colony hybridization method, and the contig to cover in between the DNA markers

at both sides as shown in Fig. 1 was completed. The cosmid clones NIT7/2 and TO3-2 comprising a part of the contig determined the base sequences by a conventional method.

Subsequently, the base sequences of the cosmid clones NIT7/2 and TO3-2 comprising a part of the contig were analyzed using "Genscan" (manufactured by Mitsubishi Space Software), by taking a parameter to Arabidopsis whose all genomes have been recently determined into consideration, because the genome DNA sequences are similar to those of radish. As a result, a promoter portion where the Rf gene seems to be expressed, a structural gene portion containing intron, and a terminator portion were discovered. In addition, the genes in the form to be translated into protein where intron is removed and the amino acid sequence to the genes were obtained.

### Embodiment 3: Sub-cloning of genomic DNA region

HpaI-SwaI fragments of DNA comprising 1 to 8,553 base sequences described in SEQ ID NO 1 sufficiently containing the terminator from the promoter assumed by "Genscan" were separated from the vector by gel electrophoresis using agarose for fragment collection (manufactured by FMC). A gel containing DNA fragments was digested by the gel degrading enzyme (manufactured by Epicentre Technologies), and DNA was collected. In addition, cloning fragments where the obtained fragments were cut by a restriction enzyme BamHI were obtained. These DNA fragments were sub-cloned to a pGEM-T easy vector (manufactured by Promega), and cds6BT/pGEM-T easy [vector] was obtained. Hereafter, details will be described.

Addition of 1 µg of NIT7/2 cosmid DNA and 10 units of restriction enzyme HpaI (manufactured by Takara Shuzo) into 100 µl of 1×K restriction enzyme buffer solution (20 mM Tris-HCl (pH8.5), 10 mM MgCl<sub>2</sub>, 1 mM

dithiothreitol and 100 mM KCl) was conducted, and the mixture was heated for 1 hour at 37 degrees C.

After heating, after 10  $\mu$ l of 3M sodium acetate (pH 5.6) and 250  $\mu$ l of ethanol were added and agitated, the mixture was cooled down at -80 degrees C for 5 minutes, and this was centrifuged at 15,000 rpm and 4 degrees C for 15 minutes. A supernatant is removed, and the precipitation was dried for 5 minutes using a centrifugal vacuum drier. Addition of 89  $\mu$ l of sterilized water into the collected DNA precipitation was conducted and the precipitation was dissolved.

Addition of 10  $\mu$ l 10 $\times$ H restriction enzyme buffer solution (500 mM Tris-HCl (pH 7.5), 100 mM  $MgCl_2$ , 10 mM dithiothreitol and 1000 mM NaCl), 1  $\mu$ l of 10 unit/ $\mu$ l restriction enzyme *Swa*I (manufactured by Takara Shuzo) into the dissolved DNA solution was conducted, and [the mixture] was heated at 25 degrees C for 1 hour. Addition of 11  $\mu$ l of 10 $\times$  loading buffer solution (1 % SDS, 50 % Glycetrol and 0.05 % Bromophenol Blue) was conducted.

After mixing 1.2 g of low melting-point agarose, SeaPlaqueGTG agarose (manufactured by FMC) and 150 ml of 1  $\times$  TAE (40 mM Tris-acetate, 1 mM EDTA) buffer solution, [the mixture] was heated to 100 degrees C and agarose was dissolved, and it was cooled down to 45 degrees C while agitating. A comb with 30 mm of width  $\times$  1 mm of thickness was placed on a gel tray of 14 cm  $\times$  15 cm, and the cooled gel was poured into the gel tray and the gel was hardened. The DNA where loading dye was added was poured into the gel comb, and electrophoresis was conducted at 1  $\times$  TAE and 30 V/30 cm of voltage for 18 hours.

The electrophoresis gel was transferred into a 0.5  $\mu$ g/ml ethidium bromide/ 1  $\times$  TAE solution, and dyeing was conducted. The gel was placed on a transilluminator where 365 nm of long-wave ultraviolet rays was irradiated, and target 8,546 bp of fragments were cut using a sterilized surgical knife. In addition, the gel was chopped to be approximately 1 mm square of

fragments, and transferred into pre-weighed 2 ml microtube, and the gel was weighed.

Addition of 1  $\mu$ l of 50x GELase buffer (2M Bis-Tris (pH6.0), 2M NaCl) into 50 mg of gel was conducted. The tube containing the gel was placed into a dry heat block heated to 68 degrees C; the tube was sometimes agitated up and down; the gel was heated for 10 minutes and was completely dissolved. This tube was transferred to a dry heat block at 45 degrees C, and the tube was sometimes agitated up and down, and the gel was heated for 5 minutes. Addition of 1 unit of GELase (manufactured by Epicentre Technologies) into 200 mg of gel weight in this tube was conducted, and the tube was sometimes agitated up and down and the gel was heated for 30 minutes.

Addition of 1/3 of gel volume of 10M ammonium acetate (pH 7.0) was conducted and [the tube] was agitated, and centrifuge was conducted at 15,000 rpm for 5 minutes. The supernatant was transferred to a new 2-ml microtube, and 2 volumes of ethanol were added to the supernatant. After the tube was agitated, it was centrifuged at 15,000 rpm and 4 degrees C for 20 minutes. The supernatant was removed, and 1 ml of 70 % ethanol was further added quietly, and it was centrifuged at 15,000 rpm and 4 degrees C for 5 minutes. The supernatant was removed, and the precipitation was dried for 5 minutes using a centrifugal vacuum drier. Addition of 20  $\mu$ l of TE buffer solution (10 mM Tris-HCl (pH 8.0), 1 mM EDTA) into the precipitation was conducted, and the precipitation was completely dissolved, and the DNA fragments were collected.

Addition of 10  $\mu$ l of 10xK restriction enzyme buffer solution (200 mM Tris-HCl (pH 8.5), 100 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, 1,000 mM KCl), 68  $\mu$ l of dH<sub>2</sub>O, 2  $\mu$ l of 10 unit/ $\mu$ l restriction enzyme BamHI (manufactured by Takara Shuzo) into the collected 20  $\mu$ l of DNA solution was conducted, and the mixture was heated at 30 degrees C for 1 hour. After heating, 10  $\mu$ l of 3M sodium acetate (pH 5.6) and 250  $\mu$ l of ethanol were added and agitated, and then, the mixture was cooled at -80 degrees C for 5 minutes and centrifuged

at 15,000 rpm and 4 degrees C for 15 minutes. The supernatant was removed, and 1 ml of 70 % ethanol was further added quietly, and then, the mixture was centrifuged at 15,000 rpm and 4 degrees C for 5 minutes. The supernatant was removed, and the precipitation was dried for 5 minutes using the centrifugal vacuum drier. Addition of 20  $\mu$ l of sterilized water to the collected DNA precipitation, and the precipitation was dissolved. After 55  $\mu$ l of sterilized water, 10  $\mu$ l of 10 % PCR buffer solution (100 mM Tris-HCl (pH 8.3), 500mM KCl), 6  $\mu$ l of 25 mM MgCl<sub>2</sub>, 8  $\mu$ l of 2.5 mM dNTP mix and 1  $\mu$ l of 5 unit/ $\mu$ l rTaq DNA polymerase (manufactured by Takara Shuzo) and the mixture was mixed, it was heated at 72 degrees C for 30 minutes and dATP was added to 3' terminal.

The reaction liquid was transferred to an ultrafiltration filter unit: Microcon-50 (manufactured by Millipore), and the reaction liquid was centrifuged at 5,000 rpm and 4 degrees C for 20 minutes. Water in the trap was disposed and 100  $\mu$ l of sterilized water was added, and the mixture was centrifuged at 5,000 rpm and 4 degrees C for 20 minutes. Addition of 20  $\mu$ l of TE buffer solution (10 mM Tris-HCl (pH 8.0) and 1 mM EDTA) was conducted, and the filter unit was removed and the posture was changed upside down, and it was mounted to a new microtube. The microtube was centrifuged at 3,000 rpm and 4 degrees C for 5 minutes.

After 1  $\mu$ l of 50 ng/ $\mu$ l pGEM-T easy vector (manufactured by Promega) and 6  $\mu$ l of DNA Ligation Kit Ver.2 (manufactured by Takara Shuzo) solution I were mixed into 5  $\mu$ l of purified DNA obtained by the above-mentioned method, the mixture was incubated at 16 degrees C for 30 minutes.

The reaction liquid was transferred to the ultrafiltration filter unit: Microcon-50 (manufactured by Millipore) along with 100  $\mu$ l of sterilized water, and the filter unit was centrifuged at 5,000 rpm and 4 degrees C for 20 minutes. Water in the trap was disposed and 100  $\mu$ l of sterilized water was added again, and the filter unit was centrifuged at 5,000 rpm and 4 degrees C for 20 minutes. The filter unit was removed and the posture was changed

upside down, and it was mounted to a new microtube. The microtube was centrifuged at 3,000 rpm and 4 degrees C for 5 minutes, and DNA in the filter unit was collected.

The collected DNA was set up together with the tube on the ice and cooled. Placement of 30  $\mu$ l of *Escherichia coli* DH10B for electroporation (manufactured by Gibco BRL) into a tube was conducted, and the tube was slightly mixed. The *Escherichia coli* mixed with DNA was transferred to a pre-iced cuvette for electroporation (1 mm of electrode intervals) (manufactured by USA Scientific Plastics). Electroporation was conducted under 1.25 kv, 129  $\Omega$  and 50  $\mu$ F of conditions using Electro Cell Manipulator 600 (manufactured by BTX), and immediately after that, 500  $\mu$ l of SOC culture medium (manufactured by Gibco BRL) warmed to 37 degrees C was added to the cuvette. The *Escherichia coli* were transferred to a 10-ml culturing tube, and it was shaken & cultured at 37 degrees C for 1 hour. The cultured *Escherichia coli* were spread to a LB agar medium (1 % Bacto-Tryptone, 0.5 % Bacto-Yeast Extract, 1 % NaCl and 1.5 % Bacto-Agar) where 100  $\mu$ g/ml of Ampiciline (manufactured by Waco Pure Chemical Industries), 20  $\mu$ g/ml X-Gal (manufactured by Takara Shuzo) and 1 mM IPTG (manufactured by Takara Shuzo), and it was cultured at 37 degrees C for 18 hours or longer.

A white colony appeared on the agar medium and it was cultured on 2 ml of the LB culture medium where 100  $\mu$ g/ml of Ampicillin was added at 37 degrees C for 18 hours or longer. Plasmid DNA was extracted from the cultured *Escherichia coli* using a usual method. It was confirmed that target fragments were cloned in the plasmid DNA via cutting with a restriction enzyme EcoRI (manufactured by Takara Shuzo), and *cds6BT/pGEM-T easy* [vectors] were obtained.

Each of *Escherichia coli* DH10B where *cds6BT/pGEM-T easy* [vectors] obtained using the above-mentioned method was maintained was cultured on 100 ml of the LB agar medium where 100  $\mu$ g/ml Ampicillin was added at 37

degrees C for 18 hours. According to the alkali SDS method, the culture was purified using a Qiagen Midi kit (manufactured by Qiagen).

#### Embodiment 4-1: Production of vector for plant transformation (1)

After *cds6BT/pGEM-Teasy* [vectors] were cut by the restriction enzyme *EcoRI*, they were separated from a vector by the gel electrophoresis using agarose for fragment collection, and the collected DNA fragments were cloned to the *EcoRI* section of vector *pKM424* for plant transformation (a vector where fragments of *CaMV35S* promoter: *GUS* gene: *NOS* terminator were added is *pLAN421* (Plant Cell Reports 10 (1991) 286-290 vector), and a vector for plant transformation *cds6BT/pKM424* were obtained. Hereafter, details will be described.

Addition of 1 µg of *cds6BT/pGEM-T easy* DNA and 10 units of restriction *EcoRI* (manufactured by Takara Shuzo) into 100 µl of 1×H restriction enzyme buffer solution (50 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol and 100 mM NaCl) was conducted, and the mixture was heated at 37 degrees C for 1 hour.

Hereafter, the *EcoRI* fragments containing *cds6BT* were separated and collected from *cds6BT/pGEM-T easy* using the similar method where the *HpaI*-*SwaI* fragments were extracted.

Addition of 1 µg of vector for plant transformation *pKM424* and 10 units of restriction enzyme *EcoRI* (manufactured by Takara Shuzo) into 100 µl of 1×H restriction enzyme buffer solution (50 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol and 100 mM NaCl) was conducted, and the mixture was heated at 37 degrees C for 1 hour. After heating, 100 µl of 1M Tris-HCl (pH 8.0) and 1 unit of Bacterial Alkaline Phosphatase (manufactured by Takara Shuzo) were added and mixed, and then, the mixture was heated at 50 degrees C for 1 hour and dephosphorylated.

Phenol chloroform saturated with 200 µl of TE buffer solution (10 mM Tris-HCl (pH 8.0) and 1 mM EDTA) was added, and the mixture was

intensely agitated. After centrifuging at 15,000 rpm for 5 minutes, the supernatant was transferred to a new tube. The similar operation was repeated once more, and a protein was removed. After 20  $\mu$ l of 3M sodium acetate (pH 5.6) and 500  $\mu$ l of ethanol were added and agitated, the mixture was cooled at -80 degrees C for 5 minutes and centrifuged at 15,000 rpm and 4 degrees C for 15 minutes. The supernatant was removed, and 1 ml of 70 % ethanol was further added, and the mixture was centrifuged at 15,000 rpm and 4 degrees C for 5 minutes. The supernatant was removed, and the precipitation was dried for 5 minutes using a centrifugal vacuum drier. Addition of 100  $\mu$ l of TE buffer solution (10 mM Tris-HCl (pH 8.0) and 1 mM EDTA) into the precipitation was conducted and the precipitation was completely dissolved, and the concentration was adjusted to 10 ng/ $\mu$ l.

After 10  $\mu$ l of purified EcoRI fragments, 1  $\mu$ l of dephosphorylated pKM424 vector and 11  $\mu$ l of DNA ligation Kit Ver.2 (manufactured by Takara Shuzo) solution were mixed, the mixture was incubated at 16 degrees C for 30 minutes.

The above-mentioned reaction liquid was transferred to the ultrafiltration filter unit: Microcon-50 (manufactured by Millipore) along with 100  $\mu$ l of sterilized water, and the mixture was centrifuged at 5,000 rpm and 4 degrees C for 20 minutes. Water in a trap was disposed and 100  $\mu$ l of sterilized water was added again, and the mixture was centrifuged at 5,000 rpm and 4 degrees C for 20 minutes. The filter unit was removed, and the posture was changed upside down and mounted to a new microtube.

The mixture was centrifuged at 3,000 rpm and 4 degrees C for 5 minutes, and DNA in the filter unit was collected.

The collected DNA was set up together with the tube on the ice and cooled. Placement of 30  $\mu$ l of *Escherichia coli* DH10B for electroporation (manufactured by Gibco BRL) into a tube was conducted, and the tube was slightly mixed. The *Escherichia coli* mixed with DNA were transferred to a pre-iced cuvette for electroporation (1 mm of electrode intervals)

(manufactured by USA Scientific Plastics). Electroporation was conducted under 1.25 kv, 129  $\Omega$  and 50  $\mu$ F of conditions using Electro Cell Manipulator 600 (manufactured by BTX), and immediately after that, 500  $\mu$ l of SOC culture medium (manufactured by Gibco BRL) warmed to 37 degrees C was added to the cuvette. The *Escherichia coli* were transferred to a 10-ml culturing tube, and it was shaken & cultured at 37 degrees C for 1 hour. The cultured *Escherichia coli* was spread to a LB agar medium (1 % Bacto-Tryptone, 0.5 % Bacto-Yeast Extract, 1 % NaCl and 1.5 % Bacto-Agar) where 50  $\mu$ g/ml of Spectinomycin (manufactured by Sigma) was added, and it was cultured at 37 degrees C for 18 hours or longer.

A white colony appeared on the agar medium and it was cultured on 2 ml of the LB culture medium where 50  $\mu$ g/ml of Spectinomycin was added at 37 degrees C for 18 hours or longer. Plasmid DNA was extracted from the cultured *Escherichia coli* using a usual method. It was confirmed that target fragments were cloned in the plasmid DNA via cutting with a restriction enzyme HindIII (manufactured by Takara Shuzo), and *cds6BT/pKM424* were obtained.

*Escherichia coli* DH10B where *cds6BT/pKM424* were maintained using the above-mentioned method were cultured on 250 ml of the LB agar medium where 50  $\mu$ g/ml Spectinomycin was added at 37 degrees C for 18 hours. According to the alkali SDS method, the culture was purified using a Qiagen Midi kit (manufactured by Qiagen).

#### Embodiment 4-2: Production of vector for plant transformation (2)

After lambda clone CHI (see Fig. 2, cloning fragment length: approximately 17 kb) where the base sequences described in SEQ ID NO 1 were sufficiently maintained was cut by the restriction enzyme NotI (manufactured by Takara Shuzo) existing in the multiple cloning sections, it was separated from a vector by the gel electrophoresis using agarose for fragment collection, and the collected DNA fragments were cloned to the NotI

section of vector pBGRZ2 for plant transformation (Bioscience and Industry 55 (1997) 37-39), and a vector for plant transformation CHI/pBGRZ2 was obtained. Hereafter, details will be described.

Addition of 1 µg of lambda clone CHI DNA and 10 units of restriction NotI (manufactured by Takara Shuzo) into 100 µl of 1×H restriction enzyme buffer solution (50 mM Tris-HCl (pH 7.5), 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol, 100 mM NaCl, 0.01 % BSA and 0.01 % TritonX-100) was conducted, and the mixture was heated at 37 degrees C for 1 hour. Hereafter, the NotI fragments of the lambda clone CHI were separated and collected using the similar method where the HpaI-SwaI fragments were extracted.

Addition of 1 µg of vector for plant transformation pBGRZ2 and 10 units of restriction enzyme NotI (manufactured by Takara Shuzo) into 100 µl of 1×H restriction enzyme buffer solution (50 mM Tris-HCl (pH7.5), 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol, 100 mM NaCl, 0.01 % BSA and 0.01 % TritonX-100) was conducted, and the mixture was heated at 37 degrees C for 1 hour. After heating, 100 µl of 1M Tris-HCl (pH 8.0) and 1 unit of Bacterial Alkaline Phosphatase (manufactured by Takara Shuzo) were added and mixed, and then, the mixture was heated at 50 degrees C for 1 hour and dephosphorylated.

Phenol chloroform saturated with 200 µl of TE buffer solution (10 mM Tris-HCl (pH 8.0) and 1 mM EDTA) was added, and the mixture was intensely agitated. After centrifuging at 15,000 rpm for 5 minutes, the supernatant was transferred to a new tube. The similar operation was repeated once more, and a protein was removed. After 20 µl of 3M sodium acetate (pH 5.6) and 500 µl of ethanol were added and agitated, the mixture was cooled at -80 degrees C for 5 minutes and centrifuged at 15,000 rpm and 4 degrees for 15 minutes. The supernatant was removed, and 1 ml of 70 % ethanol was further added, and the mixture was centrifuged at 15,000 rpm and 4 degrees C for 5 minutes. The supernatant was removed, and the precipitation was dried for 5 minutes using a centrifugal vacuum drier.

Addition of 100  $\mu$ l of TE buffer solution (10 mM Tris-HCl (pH8.0) and 1 mM EDTA) to the precipitation was conducted and the precipitation was completely dissolved, and the concentration was adjusted to 10 ng/ $\mu$ l.

After 10  $\mu$ l of purified NotI fragments, 1  $\mu$ l of dephosphorylated pBIGRZ2 vector and 11  $\mu$ l of DNA ligation Kit Ver.2 (manufactured by Takara Shuzo) solution I were mixed, the mixture was incubated at 16 degrees C for 30 minutes.

The above-mentioned reaction liquid was transferred to the ultrafiltration filter unit: Microcon-50 (manufactured by Millipore) along with 100  $\mu$ l of sterilized water, and the mixture was centrifuged at 5,000 rpm and 4 degrees C for 20 minutes. Water in a trap was disposed and 100  $\mu$ l of sterilized water was added again, and the mixture was centrifuged at 5,000 rpm and 4 degrees C for 20 minutes. The filter unit was removed, and the posture was changed upside down and mounted to a new microtube. The mixture was centrifuged at 3,000 rpm and 4 degrees C for 5 minutes, and DNA in the filter unit was collected.

The collected DNA was set up together with the tube on the ice and cooled. Placement of 30  $\mu$ l of *Escherichia coli* DH10B for electroporation (manufactured by Gibco BRL) into a tube was conducted, and the tube was slightly mixed. The *Escherichia coli* mixed with DNA were transferred to a pre-iced cuvette for electroporation (1 mm of electrode intervals) (manufactured by USA Scientific Plastics). Electroporation was conducted under 1.25 kv, 129  $\Omega$  and 50  $\mu$ F of conditions using Electro Cell Manipulator 600 (manufactured by BTX), and immediately after that, 500  $\mu$ l of SOC culture medium (manufactured by Gibco BRL) warmed to 37 degrees C was added to the cuvette. The *Escherichia coli* were transferred to a 10-ml culturing tube, and it was shaken & cultured at 37 degrees C for 1 hour. The cultured *Escherichia coli* were spread to a LB agar medium (1 % Bacto-Tryptone, 0.5 % Bacto-Yeast Extract, 1 % NaCl, 1.5 % Bacto-Agar) where 25

µg/ml of Kanamycin (manufactured by Wako Pure Chemical Industries) was added, and they were cultured at 37 degrees C for 18 hours or longer.

A colony appeared on the agar medium and it was cultured on 2 ml of the LB culture medium where 25 µg/ml of kanamycin was added at 37 degrees C for 18 hours or longer. Plasmid DNA was extracted from the cultured *Escherichia coli* using a usual method. It was confirmed that target fragments were cloned in the plasmid DNA via cutting with a restriction enzyme HindIII (manufactured by Takara Shuzo), and CHI/pBGRZ2 was obtained.

*Escherichia coli* DH10B where CHI/pBGRZ2 obtained using the above-mentioned method were maintained were cultured on 250 ml of the LB agar medium where 25 µg/ml kanamycin was added at 37 degrees C for 18 hours. According to the alkali SDS method, the culture was purified using a Qiagen Midi kit (manufactured by Qiagen).

Embodiment 5: Introduction of *Agrobacterium* to vector for plant transformation

Competent cells of *Agrobacterium* were prepared, and the *cds6BT/pKM424* vectors and the CHI/pBGRZ2 vectors obtained in Embodiments 4-1 and 4-2 were introduced into the prepared *Agrobacterium* EHA101 for plant transformation, respectively. Hereafter, the details will be described.

The competent cells for electroporation of *Agrobacterium* EHA101 were prepared using the method mentioned below. *Agrobacterium* EHA101 was streaked over the LB agar culture medium where 50 µg/ml kanamycin (manufactured by Wako Pure Chemical Industries) and 25 µg/ml Chloramphenicol (manufactured by Wako Pure Chemical Industries), and [*Agrobacterium* EHA101] was cultured at 28 degrees for 24 hours or longer, and a single colony was obtained. The LB culture medium where 50 µg/ml kanamycin and 25 µg/ml Chloramphenicol were added was placed into a 50-

ml centrifuging tube and a colony with approximately 1 mm of diameter was inoculated, and the centrifuging tube was shaken and cultured at 28 degrees C for 40 hours. The culture solution was centrifuged at  $1,500\times g$  and 4 degrees C, and harvest was conducted. Placement of 40 ml of iced sterilized 10 % glycerol to the tube where the supernatant was removed was conducted and a fungus body was re-suspended, and the tube was centrifuged and harvested at  $1,500 \times g$  and 4 degrees C. This operation was repeated twice. Addition of 500  $\mu$ l of iced sterilized 10 % glycerol into the obtained fungus body was conducted, and the mixture was re-suspended. Divided injection of 100  $\mu$ l each of fungus body into the sterilized microtubes was conducted, and after these were frozen with liquid nitrogen, they were stored in an 80-degrees C freezer.

The competent cells for electroporation of *Agrobacterium* EHA101 were melted on iced water. Placement of 40  $\mu$ l of electrocompetent cells into pre-cooled 1.5 ml tube was conducted, and 100 ng of *cds6BT/pKM424* or *CHI/pBIGRZ2* plasmid DNA was added and slightly mixed.

*Agrobacterium* mixed with DNA was transferred to a pre-iced cuvette (manufactured by USA Scientific Plastics) for electroporation (electrode intervals: 1 mm). Electroporation was conducted under conditions of 1.44 kv, 129  $\Omega$  and 50  $\mu$ F using Electro Cell Manipulator 600 (manufactured by BTX), and immediately after that, 500  $\mu$ l of SOC culture medium (manufactured by Gibco BRL) was added. *Agrobacterium* was transferred to a 10-ml culturing tube, and the tube was shaken and cultured at 30 degrees C for 1 hour.

Cultured *Agrobacterium* where a *cds6BT/pKM424* were introduced was spread over the LB agar culture medium (1 % Bacto-Tryptone, 0.5 % Bacto-Yeast Extract, 1 % NaCl and 1.5 % Bacto-Agar) where 50  $\mu$ g/ml kanamycin (manufactured by Wako Pure Chemical Industries), 25  $\mu$ g/ml Chloramphenicol (manufactured by Wako Pure Chemical Industries), 50  $\mu$ g/ml Spectinomycin (manufactured by Sigma) and 2.5 $\mu$ g/ml Tetracycline

(manufactured by Sigma) were added, and it was cultured at 28 degrees C for 24 hours or longer.

Cultured *Agrobacterium* where a CHI/pBGRZ2 vectors were introduced was spread over the LB agar culture medium where 50 µg/ml kanamycin (manufactured by Wako Pure Chemical Industries), 25 µg/ml Chloramphenicol (manufactured by Wako Pure Chemical Industries) and 30 µg/ml Hygromycin (manufactured by Sigma) were added, and it was cultured at 28 degrees C for 24 hours or longer.

A colony appeared on the agar culture medium and it was cultured on 2 ml of LB culture medium where antibiotics adapted to each vector were added at 30 degrees C for 24 hours or longer. Plasmid DNA was extracted from the cultured *Agrobacterium* using a usual method, and it was confirmed that *cds6BT/pKM424* vectors or *CHI/pBGRZ2* vectors were introduced by cutting with the restriction enzyme *HindIII* (manufactured by Takara Shuzo). For the confirmed clones, after an equal amount of sterilized 80 % glycerol was added to the culture solution, which was cultured for 24 hours, and they were mixed; they were stored at -80 degrees C and used for rapeseed transformation.

#### Embodiment 6: Production of rapeseed transformant

The transformation to rapeseed was conducted as follows: CMS rapeseed (SW18) seeds having radish-derived *cms* causal genes *ORF125* were sterilized with a 10 % hydrochlorous acid solution, and they were sprouted on the MS culture medium (T. Murashige and F. Skoog *Physiol. Plant.* 15:485, 1962) not containing hormone. Only hypocotyl parts were cut off from seedling where 7 days to 14 days passed after sprout, they were cut into 3 to 5 mm of length, and they were pre-cultured on the MS culture medium (M5519 provided by Sigma) + 3 % sucrose + 1 mg/l 2,4-D and 0.4 % agarose (Sigma, type I) at 23 degrees C for 12 to 16 hours. At this time, for the

protection culture, cocultivation with tobacco-derived cell line BY-2 was conducted.

In the meantime, *Agrobacterium* containing CHI/pBGRZ2 was cultured at 28 degrees C for 8 to 48 hours, and was proliferated to approximately  $OD_{600} = 1.0$ . The fungus body of *Agrobacterium* was suspended in a liquid MS hormone-free culture medium. The cut hypocotyl and this *Agrobacterium* solution were mixed, and the mixture was cocultivated for approximately 20 minutes. After the cocultivation, the hypocotyl where *Agrobacterium* was removed with a paper filter was cultured, for example, in a culture medium with MS basic culture medium + B5 vitamin (Sigma, M0404) + 3 % sucrose + 1 mg/l 2,4-D, for 2 days and infected. After the infection, the hypocotyl was transplanted to a sterile culture medium where an antibiotic: carbenicillin (geopen by Pfizer or carbenicillin disodium salts by GIBCO-BRL) was added with 500 mg/l of concentration into MS basic culture medium + B5 vitamin + 3 % sucrose + 1 mg/l 2,4-D, and *Agrobacterium* was removed.

After five days to one week were passed with the above-mentioned sterile culture medium, the hypocotyl was added to MS basic culture medium + B5 vitamin + 1 % sucrose + 3mg/l benzylaminopurine + 500mg/l carvenicillin, and it was cultured on a culture medium where 5 mg/l silver nitrate and 5 to 30 mg/l kanamycin for selection (kanamycin sulfate salt, Nacalai Tesque) were added, for 14 days to 21 days. At this time, since green callus might appear, they were immediately transplanted to a culture medium in the next step.

As the culture medium in the next step, for example, there is a selection medium containing MS culture medium (Sigma, M5519) + 1 % sucrose + 3 mg/l benzylaminopurine + 1 mg/l zeatin + 500 mg/l carbenicillin + 5 to 30 mg/l kanamycin. The hypocotyl where the callus was formed from the incision was transplanted to this culture medium and cultured at 23 degrees

C for 3 weeks, and then, 3 to 5 times of transplants were repeated every 3 weeks until the green callus appeared.

The green callus was cut off from the hypocotyl immediately as soon as it was discovered, and it was transferred to the culture medium with the same composition. After that, when only green portions were cut off and joined, adventitious buds were formed with 1 % to 30 % of probability. After the adventitious buds were transferred to B5 basic culture medium (Sigma, G5893) + 3 % sucrose + 1 mg/l benzylaminopurine and cultivated thereafter, they were rooted in the culture medium containing MS culture medium (Sigma, M5519) + 3 % sucrose + 0.1 mg/l naphthalene acid + 0.1 mg/l benzylaminopurine.

#### Embodiment 7: Analysis of transformant (detection of introduced DNA)

One leaf was extracted from one transformation individual where buds were formed obtained in Embodiment 6, and DNA was isolated using a DNA isolation kit (DNeasy plant mini) by QIAGEN.

Three sections (sections *a*, *b* and *c*) of the introduced DNA fragments were detected using the PCR method (results are shown in Fig. 3). The section *a* contains 568bp from 3,186bp to 3,753bp of base sequences in SEQ ID NO 1, and [5'-GAAGCAAAAAAGAAAACGAGCAGAG-3'] (SEQ ID NO 4) and [5'-CCAAAAATCCGAAATCCGAATAGAC-3'] (SEQ ID NO 5) were used as a forward primer and a reverse primer, respectively. The section *b* contains 244bp from 4,869bp to 5,112bp of base sequences in SEQ ID NO 1, and [5'-CTCGGCTCTGGGTTTAGTGA-3'] (SEQ ID NO 6) and [5'-TCCACAAACCCTAGCCAACA-3'] (SEQ ID NO 7) were used as a forward primer and a reverse primer, respectively. The section *c* contains 485bp from 7,766bp to 8,250bp of base sequences in SEQ ID NO 1, and [5'-GCTTATGCTTCTCTGGTTCGCCTC-3'] (SEQ ID NO 8) and [5'-CTCAGTTTTCGTACCTTACACAATGC-3'] (SEQ ID NO 9) were used as a forward primer and a reverse primer, respectively.

After 12.1  $\mu$ l of sterilized water, 2  $\mu$ l of  $10 \times$  PCR buffer solution (100 mM Tris-HCl (pH8.3), 500 mM KCl), 1.2  $\mu$ l of 25 mM  $MgCl_2$ , 1.6  $\mu$ l of 2.5 mM dNTP mix, 1  $\mu$ l of 10  $\mu$ M forward primer solution of each section, 1  $\mu$ l of 10  $\mu$ M reverse primer solution of each section, 0.1  $\mu$ l of 5 unit/ $\mu$ l rTaq DNA polymerase (manufactured by Takara Shuzo) were added to 1  $\mu$ l of transformant DNA solution (50 ng/ $\mu$ l) and mixed, DNA was amplified by repeating a cycle of 94 degrees for 40 seconds, 55 degrees C for 30 seconds and 72 degrees for 1 minute, 33 times. For a thermal cycler, UNOII (manufactured by Biometra) was used. After the reaction was completed, an amplified product was confirmed by gel electrophoresis of 4 % Nusive 3:1 Agarose (manufactured by FMC)/  $1 \times$ TBE (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA) buffer solution (see Fig. 3).

As a result, it was ascertained that the section *a* was not introduced into this transformed rapeseed. An amplified product with the same size as the positive control was obtained at the remaining two sections (the sections *b* and *c*), respectively, and it was confirmed that the DNA was incorporated into the transformed rapeseed.

Embodiment 8: Analysis of transformant (confirmation of decrease in cms causal protein ORF125 accumulation)

One bud from the same individual in Embodiment 7, and a decrease of ORF125, which is CMS protein) was analyzed by the Western blotting method. The results are shown in Fig. 4.

(1) Protein extraction from transformed individual:

The Western blotting method, which is a protein extraction method, was conducted based upon the method by N. Koizuka et al. (Theor Appl Genet (2000) 100:949-955).

Specifically, one bud (1 mm of length) of the obtained transformed rapeseed and 100  $\mu$ l of iced protein extracted buffer solution (50 mM Tris-HCl (pH7.5), 2 % (W/V) SDS) were placed into an iced mortar, and they were

ground by a pestle. This liquid was transferred to a micro-centrifugal tube, and it was centrifuged at 15,000 rpm and 4 degrees C for 15 minutes. After centrifuging, a supernatant solution was transferred to a new micro-centrifugal tube, and it was heated at 100 degrees for 5 minutes. The supernatant solution was centrifuged at 15,000 rpm and 4 degrees C for 15 minutes again, and a supernatant solution was transferred to a new micro-centrifugal tube, and this solution was regarded as a SDS soluble protein solution. The concentration of the SDS soluble protein solution was measured using a protein quantitative kit (manufactured by Bio-rad) by a Bradford method. Along with this, the SDS soluble protein solution was similarly extracted from buds from rapeseed of a cytoplasmic male sterile line and rapeseed of fertility restored line, as well, and the concentration was measured, respectively.

(2) Isolation of protein by SDS-PAGE method and transcription to PVDF film: Western blotting

Placement of 15 µg of the SDS soluble protein per lane onto a 10 % SDS polyacrylic amide gel with 7 cm × 10 cm square was conducted, and the protein was isolated by electrophoresis. In addition, for the purpose of comparison of ORF125 protein accumulation, a diluted line of the cytoplasmic male sterile line rapeseed was similarly placed and isolated. The electrophoresis conditions were for 1 hour with 10 mA and for 1 hour with 15 mA. After the electrophoresis, the protein in the polyacrylic amide gel was transcribed onto a PVDF film (manufactured by Millipore) under the condition with 100 mA for 1 hour.

(3) Detection of protein using antibody: Western blotting

The PVDF film where the protein was transcribed was divided into upper and lower two films, and they were transferred to 10 ml of blocking solution (20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 0.05 % Tween20 and 5 % skim milk), and they were shaken for 1 hour and blocked. ATPA as a control of mitochondria protein mass and ORF125, which is a cytoplasmic male

sterility related protein, were detected from the upper PVDF film and the lower PVDF film, respectively. The PVDF films were transferred to 10 ml of primary antibody reaction liquid (100 µl of ATPA monoclonal antibody was added for ATPA detection, and rabbit antiserum to 2 µl of ORF125 was added for ORF125 detection into 10 ml of blocking solution (M.Iwabuchi et al. *Plant Molecular Biology* (1999) 39: 183-188)) was conducted, and they were shaken for 18 hours. The PVDF films were transferred to 100 ml of TTBS (20 mM Tris-HCl (pH 7.5), 500 mM NaCl and 0.05 % Tween20) and shaken for 10 minutes. This operation was repeated three times, and excess primary antibody liquid was washed out. The PVDF films were transferred to 10 ml of secondary antibody reaction liquid (goat antimouse IgG (manufactured by Amersham) where 10 µl of peroxidase was added for ATPA detection and goat antimouse IgG (manufactured by Bio-rad) where 10 µl of alkali phosphatase was added for ORF125 detection into 10 ml of blocking solution (M. Iwabuchi et al. *Plant Molecular Biology* (1999) 39:183-188)), and they were shaken for 1 hour, respectively. The PVDF films were transferred to 100 ml of TTBS (20 mM Tris-HCl (pH 7.5), 500 mM NaCl and 0.05 % Tween20), and they were shaken for 10 minutes. This operation was repeated three times, and an excess secondary antibody was washed out. A chemical emission system "ECL+" (manufactured by Amersham) to peroxidase was used for ATPA detection, and exposure detection was conducted for 5 seconds. BCIP/NBT (manufactured by MOSS Inc.), which is a chromogenic substrate of alkali phosphatase, was used for ORF125 detection, and color development was detected for 5 minutes.

As a result, the ATPA accumulation, which is a control, was hardly changed in buds of the diphyletic cytoplasmic male sterility rapeseed, [buds of] fertility restored rapeseed and buds of transformed rapeseed where the DNA was inserted into the cytoplasmic male sterile line; however, the accumulation of ORF125 protein was decreased in the transformed rapeseed dominantly. The degree of this decrease is the same level as the fertility

restored line where the fertility restoring gene was introduced due to the cytoplasmic male sterile line by breeding (Fig. 4 and M. Iwabuchi et al. Plant Molecular Biology (1999) 39:183-188). Further, comparing the degree of decrease in ORF125 protein accumulation with the dilute line, it was ascertained that the decrease was 1/8 to 1/16 with the fertility restored rapeseed and approximately 1/8 with transformed rapeseed. As described above, since the fertility restoration in rapeseed and the decrease in ORF125 protein accumulation were strongly linked and in the agreement relationship, it was proved that the DNA sequence functions to decrease the ORF125 protein accumulation within mitochondria, and is a genome DNA sequence where the fertility restoring genes are retained.

In addition, anthers were extracted from a bloomed body and they were observed with a microscope, and it was confirmed that normal pollen was formed (Fig. 5).

#### Embodiment 9: cDNA isolation

A F<sub>2</sub> individuals, which have the Rf1 gene in homo and has pollen fertility, were selected from the F<sub>2</sub> aggregation, which was used at the time of production of gene mal, as RNA donor, and after mRNA was purified from buds, cDNA was synthesized, and cDNA was isolated using 5'-RACE or 3'-RACE method.

##### (Purification of mRNA)

All RNAs were extracted from buds of F<sub>2</sub> individual, which has the Rf1 gene having allele and has pollen fertility, by a guanidium thiocyanate method, which is a usual method, using the RNeasy kit (by Qiagen). Poly A + RNA were purified from all RNAs using mRNA Purification kit (by Amersham Pharmacia) where origo (dT) cellulose columns, and mRNA was obtained.

##### (Isolation of cDNA by 5'-RACE and 3'-RACE)

Using 1 µg of purified mRNA, cDNA was isolated using “Marathon RACE system 5'RACE 3'RACE” kit by 5'-RACE and 3'-RACE methods. As gene specific primers, 5'-GATTCCTTTCTCTTGCATTTCAG-3' (SEQ ID NO 10) and 5'-ATCTCGTCCTTTACCTTCTGTGG-3' (SEQ ID NO 11) were used for 5'-RACE and 3'-RACE, respectively. After base sequences of the obtained clone were defined, cDNA sequences were obtained (SEQ ID NO 2).

Embodiment 10: Conversion of cDNA to amino acid sequence and analysis  
(1) Using a normal genetic code, cDNA was converted into the amino acid sequences using gene analysis software “Genetyx-SV” (Software Development Co., Ltd.), and amino acid sequences described in SEQ ID NO 3 were obtained. PPR motif was analyzed using a program in Protein families' database of alignments and HMNs (hereafter, abbreviated as Pfam: <http://www.sanger.ac.uk/Software/Pfam/search.shtml>). As a result of analysis, it was discovered that a translated product of fertility restoring genes shown with SEQ ID NO 1 was a protein having 16 PPR motifs. Further, the PPR motifs were three PPR clusters. The three clusters were as follows:

- ① PPR cluster #1: PPR cluster where the 1st PPR motif from the N-terminal to the 5th PPR motif comprises successive 175 residues;
- ② PPR cluster #2: PPR cluster where the 6th PPR motif from the N-terminal to the 12th PPR motif comprises successive 245 residues;
- ③ PPR cluster #3: PPR cluster where the 13th PPR motif from the N-terminal to the 16th PPR motif comprises successive 140 residues;

Embodiment 11: Analysis of protein of the present invention

An experiment about whether or not translation inhibition occurred in *Escherichia coli* combining a transcriptional product (mRNA) of ORF125, which is a causal protein occurring Kosena cytoplasmic male sterility was conducted.

The fertility restoring genes shown in SEQ ID NO 2 were introduced into a BamHI-SphI section of an *Escherichia coli* expression vector pQE-80L (by Qiagen), and six histidine residues structured an expression vector to be added to (6XHis) N-terminal (pQEB1/cds6). Further, DNA was amplified using a primer to introduce BamHI section: CGGGATCCGCTCACAATT (SEQ ID NO 12) and an M13 primer RV (manufactured by Takara Shuzo), using pSTV29 (manufactured by Takara Shuzo) vector as a mold. For the DNA amplification, Takara LA PCR Kit (manufactured by Takara Shuzo) was used. After the amplified DNA was cut by restriction enzymes BamHI and EcoRI (manufactured by Takara Shuzo), they were purified using Suprec-02 (manufactured by Takara Shuzo). In order to synthesize DNA fragments having BamHI and EcoRI sections in the 5'-UTR region of ORF125 gene and at the both sides of 25 amino acids of ORF125, PCR was conducted using two primers in accordance with Fujimoto method (PCR experimental protocol of plants: Actuality of synthesized DNA, pp. 84-87 (Shujunsha). For the primers,

125-5'BamHI:

GCGGATCCCAATTTTCATTCTGCATCACTCTCCCTGTCGTTATCGACCTCG  
CAAGGTTTTTTGAAACGGCCGAAACGGGAAGTGACAATACCGCTTTTCTT  
C (SEQ ID NO 13) and

125-5'EcoRI:

GGAATTCATACTTTTACATTTCAGTAGGAGTGAGATTATGACAAAAAGTG  
GACAATTTTTTCGAAAAAGGTAATCATGCATTTATATGCTGAAGAAAAGCG  
(SEQ ID NO 14) were used. The amplified DNA was cut by the restriction enzymes BamHI and EcoRI (manufactured by Takara Shuzo), DNA was purified using Suprec-02 (manufactured by Takara Shuzo). The purified DNA was joined using TaKaRa Ligation kit (manufactured by Takara Shuzo), and was transformed to *Escherichia coli* DH10B (manufactured by Gibco BRL). They were cultured in the LB agar culture medium (1 % Bacto-Tryptone, 0.5 % Bacto-Yeast Extract, 1 % NaCl, 1.5 % Bacto-Agar, 0.1 mM

IPTG and 20 µg/ml X-Gal) where 50 µg/ml of chloramphenicol (manufactured by Sigma) was added, at 37 degrees for 18 hours or longer. Plasmid was extracted from a thin blue colony using a usual method, and base sequences were confirmed. As described above, a vector where the 5'-UTR region of ORF125 gene and 174bp including 25 amino acid of ORF125 (the 7th to 180th base sequences output the base sequences described in Fig. 6) were introduced between from the EcoRI section to the transcription starting point of lacZ gene was structured (pSTV125-5' #LA6). Further, simultaneously, a vector having fragments (the 7th to 183rd of the base sequences described in Fig. 6) causing the mutation to a portion equivalent to said 174bp was obtained (pSTV125-5' #LA12).

Said vectors pSTV125-5' #LA6 and #LA12 were introduced into the *Escherichia coli* (manufactured by GibcoBRL), respectively, and they were stationarily cultured in the agar culture medium where 50 µg/ml chloramphenicol (manufactured by Wako Pure Chemicals Industries), 200 µM IPTG (manufactured by Wako Pure Chemicals Industries) and 40 µg/ml X-Gal (manufactured by Takara Shuzo) were added to the LB culture medium at 37 degrees over night, and a colony was grown, and it became thin blue. In other words, even in the *Escherichia coli* where any vector was introduced, it was the introduced LacZ genes that were expressed.

In addition, in order to introduce the pSTV125-5' vector and the pQEB1/cds6 vectors into the *Escherichia coli* as similar to the above-mentioned, they were cultured using a culture medium where 50 µg/ml ampicillin was added to the above-mentioned culture medium. When pSTV125-5' #LA6 was co-existed, the colony became white; however, when pSTV125-5' #LA12 having mutation to the introduced fragments was co-existed, the colony became thin blue, and the degree of blue was the same as the case of #LA12 singularity. Furthermore, whether or not these introduced vectors fell off from the *Escherichia coli* was confirmed by extracting the vectors using a usual method after each colony was cultured.

According to the above-mentioned results, it seems that the protein expressed within the *Escherichia coli* by the pQEBl/cds6 vectors became white colony, because the expression of LacZ genes was restrained as a result of combining with mRNA of pSTV125-5' #LA6. Further, when pSTV125-5' #LA12 co-existed, pQEBl/cds6-derived protein could not be combined with mRNA because the introduced fragment section had mutation, and as a result, the LacZ gene was expressed and the color became blue.

As described above, it is assumed that the protein having the amino acid sequences described in SEQ ID NO 3 participates in mRNA of ORF125; to be more specific, the transcriptional products of at least ORF125-5'UTR region and the code region of 25 amino acid residues of ORF125, and the expression of ORF125 protein appears to be restrained.

In other words, it is assumed that since the translational products of genes of the present invention are combined with male-sterile genes and inhibit the translation within mitochondria after being transferred to the mitochondria, the accumulation of causal protein of cytoplasmic male sterility is decreased, and the cytoplasmic male sterility is restored to the fertility.

#### Embodiment 12: Isolation of Kosena rapeseed fertility restoration gene

The genomic sequence of the fertility restoring gene was obtained from the rapeseed line, which was obtained by introducing kosena radish (*Raphanus sativus* cv. Kosena) due to cell fusion, (hereafter, referred to as kosena rapeseed) using the PCR method. DNA was extracted from 1.0 g of a leaf of kosena rapeseed using the DNA isolation kit (DNeasy plant mini) provided by Qiagen. For DNA amplification, 5'-ACATAAAATCACTAGATACTTGACATGGAGGC-3' (SEQ ID NO 30), which represents sequences from the base sequence 1,027bp to 1,059bp in SEQ ID NO 1, was set as a forward primer, and 5'-AAGAGGAGGAAGATGGCATCACAGC-3' (SEQ ID NO 31), which represents

sequences from the base sequence 7,675bp to 7,651bp, was set as a reverse primer.

After 49  $\mu$ l of sterilized water, 10  $\mu$ l of 10 $\times$ LA PCR buffer solution (manufactured by Takara Shuzo), 10  $\mu$ l of 25 mM  $MgCl_2$ , 16  $\mu$ l of 2.5 mM dNTP mix, 2  $\mu$ l of 10  $\mu$ M forward primer solution, 2  $\mu$ l of 10  $\mu$ M reverse primer solution and 1  $\mu$ l of 5 unit/ $\mu$ l TaKaRa LA Taq (manufactured by Takara Shuzo) were added to 10  $\mu$ l kosena rapeseed DNA solution (50 ng/ $\mu$ l) and they were mixed, DNA was amplified by repeating 30 times with a cycle at 98 degrees C for 20 seconds and at 68 degrees C for 15 minutes. For a thermal cycler, UNOII (manufactured by Biometra) was used. After the reaction, approximately 6 kb of amplified product was purified using an ultrafiltration filter Microcon-PCR (manufactured by Millipore). For a portion equivalent to the 4,280th to the 7,585th in SEQ ID NO 1, the base sequence 3,306bp was determined (SEQ ID NO 15) using the purified amplified product as a mold. The obtained genome base sequence of kosena rapeseed and the sequence of Yuanhong radish (*Raphanus sativus* cv. Yuanhong) were compared, and DNA base substitution was discovered only in 7bp output 3,306bp, and it was ascertained that [kosena rapeseed] has high homology.

Further, a cDNA sequence of 3' portion of fertility restoring gene of kosena rapeseed was acquired by RT-PCR method, and it was confirmed that intron was spliced in the same form as Yuanhong radish (*Raphanus sativus* cv. Yuanhong). Entire RNA was extracted from buds of kosena rapeseed using AGPC (Acid Guanidium-Phenol-Chloroform) method (Shujunsha, supplementary volume of Cell Technology, Bio-experimental Illustrated; ② Basic of Gene Analysis, p. 161-166), which is a conventional method. Using SUPERSRIPTII RNaseH<sup>-</sup> Reverse Transcriptase (manufactured by Invitrogen), cDNA was synthesized with 1  $\mu$ g of entire RNA. As a specific forward primer for 3' portion of fertility restoring gene, 5'-TGGAGTAAAGAGGAACTAAAAAGGGC-3' (SEQ ID NO 23) was used, and

as a specific reverse primer for 3' portion of fertility restoring gene, 5'-CAGACAATAGACGCATAAAAGGC-3' (SEQ ID NO 33) was used. After 14.9 µl of sterilized water, 2.5 µl of 10×PCR buffer solution (manufactured by Takara Shuzo), 1.5 µl of 25 mM MgCl<sub>2</sub>, 2 µl of 2.5 mM dNTP mix, 1.5 µl of 10 µM forward primer solution, 1.5 µl of 10 µM reverse primer solution and 0.1 µl of 5 unit/µl of TaKaRa Taq (manufactured by Takara Shuzo) were added into 1 µl of kosenara rapeseed cDNA solution and they were mixed, DNA was amplified by repeating 35 times with a cycle at 94 degrees for 40 seconds, at 60 degrees for 30 seconds and at 72 degrees for 2 minutes 35 times. For a thermal cycler, UNOII (manufactured by Biometra) was used. Addition of 1 µl of pGEM-Teasy vector (manufactured by Promega), 5 µl of 2× ligation buffer solution (manufactured by Promega) and 1 µl T4 DNA ligase (manufactured by Promega) into 3 µl of amplified product was conducted, and [the mixture] was left to stand at room temperature for one hour, and combined to the vector. This vector became transformed to Escherichia coli DH5α (manufactured by Gibco BRL) and clones were obtained. For the obtained clones, base sequences were determined by a conventional method, and it was confirmed that intron of cDNA of kosenara rapeseed was spliced as similar to Yuanhong radish (*Raphanus sativus* cv. Yuanhong). In other words, it was ascertained that the base substitution at 7bp, which was discovered from the comparison of the prior genomic sequences, existed in 5,444bp to 5,814bp of SEQ ID NO 1, and it did not affect splicing.

As described above, it was ascertained that the fertility recovering gene of kosenara rapeseed expresses mRNA in a similar form as Yuanhong radish (*Raphanus sativus* cv. Yuanhong). The cDNA sequence coding for only the translated region is shown in SEQ ID NO 16. In addition, an amino acid sequence was obtained from this cDNA sequence (SEQ ID NO 17).

Embodiment 13: Isolation of fertility restoring gene partial sequence of *Raphanus raphanistrum* (hereafter, referred to as *R. raphanistrum*)

A partial sequence of *R. raphanistrum* fertility restoring gene was isolated from cDNA.

(Purification of mRNA)

Entire RNA was extracted from buds of *R. raphanistrum* using AGPC (Acid Guanidium-Phenol-Chloroform) method (Shujunsha, supplementary volume of Cell Technology, Bio-experimental Illustrated; © Basic of Gene Analysis, p. 161-166), which is a conventional method. PolyA + RNA were purified from the entire RNA using "mRNA Purification kit" (provided by Amersham Pharmacia) where an Origo(dT) cellulose column was used, and mRNA was obtained.

(Partial sequence isolation of cDNA)

From 1 µg of purified mRNA, cDNA was synthesized by using a "Marathon RACE system 5'RACE 3'RACE" kit (manufactured by Clontech). As a specific forward primer of fertility restoring gene, 5'-GATTCCTTTCTCTTGCATTTTCAG-3' (SEQ ID NO 34) was used, and as a specific reverse primer of fertility restoring gene, 5'-ATCTCGTCCTTTACCTTCTGTGG-3' (SEQ ID NO 35) was used.

After 14.4 µl of sterilized water, 2.5 µl of 10x Pyrobest PCR buffer solution (manufactured by Takara Shuzo), 2 µl of 2.5 mM dNTP mix, 2.5 µl of 10 µM forward primer solution, 2.5 µl of 10 µM reverse primer solution, 0.1 µl of 5 unit/µl Pyrobest DNA polymerase (manufactured by Takara Shuzo) were added to 1 µl *R. raphanistrum* cDNA solution and they were mixed, DNA was amplified by repeating 30 times with a cycle at 98 degrees C for 5 seconds, at 55 degrees C for 30 seconds and at 72 degrees C for 1 minute and 30 seconds. As a thermal cycler, UNOII (manufactured by Biometra) was used. After 0.1 µl of 5 unit/µl TaKaRa Taq (manufactured by Takara Shuzo) was added into the amplified DNA solution and they were mixed, the mixture was heated at 72 degrees C for 10 minutes, and adenine nucleotide was added to the 3' terminal of DNA. Addition of 1 µl of pGEM-Teasy vector (manufactured by Promega), 5 µl of 2x ligation buffer solution (manufactured

by Promega) and 1 µl of T4 DNA ligase (manufactured by Promega) into 3 µl of the amplified product was conducted, and the mixture was left to stand at room temperature for 1 hour, and this was combined to the vector. This vector was transformed to *Escherichia coli* DH5α (manufactured by Gibco BRL) and clones were obtained. After base sequences of the obtained clones were determined using a conventional method, the cDNA partial sequence was obtained (SEQ ID NO 20). In addition, the amino acid sequence was obtained from the cDNA sequence (SEQ ID NO 21).

Embodiment 14: Isolation from ogura rapeseed fertility restoration gene

The cDNA sequence of the fertility restoring gene was isolated from the rapeseed line, which was obtained by introducing the fertility restoring genes of *Ogura radish* (*Raphanus sativus* L.) due to hybridization (hereafter, referred to as ogura rapeseed), using 5'-RACE method and 3'-RACE method. (Purification of mRNA)

Entire RNA was extracted from buds of ogura rapeseed using AGPC (Acid Guanidium-Phenol-Chloroform) method (Shujunsha, supplementary volume of Cell Technology, Bio-experimental Illustrated; ② Basic of Gene Analysis, p. 161-166), which is a conventional method. PolyA + RNA were purified from the entire RNA using "mRNA Purification kit" (provided by Amersham Pharmacia) where a Origo(dT) cellulose column was used, and mRNA was obtained.

(Partial sequence isolation of cDNA)

From 1 µg of purified mRNA, cDNA was synthesized using a "Marathon RACE system 5'RACE 3'RACE" kit (manufactured by Clontech). As a specific forward primer of fertility restoring gene, 5'-GATCCATGCATTTGTCAAGG-3' (SEQ ID NO 36) was used, and as a specific reverse primer of fertility restoring gene, 5'-CATTTGTGTAGCCTCATCTAGG-3' (SEQ ID NO 37) was used.

After 14.4 µl of sterilized water, 2.5 µl of 10x Pyrobest PCR buffer solution (manufactured by Takara Shuzo), 2 µl of 2.5 mM dNTP mix, 2.5 µl of 10 µM forward primer solution, 2.5 µl of 10 µM reverse primer solution, 0.1 µl of 5 unit/µl Pyrobest DNA polymerase (manufactured by Takara Shuzo) were added into 1 µl of ogura rapeseed cDNA solution and they were mixed, DNA was amplified by repeating 30 times with a cycle at 98 degrees C for 5 seconds, at 55 degrees C for 30 seconds, and at 72 degrees for 1 minute 30 minutes. For a thermal cycler, UNOII (manufactured by Biometra) was used. After 0.1 µl of 5 unit/µl TaKaRa Taq was added to the amplified DNA solution and they were mixed, the mixture was heated at 72 degrees for 10 minutes, and adenine nucleotide was added to 3' terminal of DNA. Addition of 1 µl of pGEM-Teasy vector (manufactured by Promega), 5 µl of 2x ligation buffer solution (manufactured by Promega), 1 µl of T4 DNA ligase (manufactured by Promega) into 3µl of the amplified product was conducted, and the mixture was left to stand at room temperature for one hour and this was combined to the vector. This vector was transformed to Escherichia coli DH5α (manufactured by Gibco BRL), and clones were obtained. After base sequences of the obtained clones were determined using a usual method, four types of cDNA partial sequences were obtained. Primers for 5'-RACE and 3'-RACE were set to a common portion to the four types based upon the obtained four types of sequence information, and cDNA was isolated, respectively.

(Isolation of cDNA by 5'-RACE and 3'-RACE)

The cDNA was synthesized as similar to the above-mentioned, and the 5'-RACE and 3'-RACE methods were conducted using a "Marathon RACE system 5'RACE 3'RACE" kit (manufactured by Clontech), and cDNA was isolated. As a gene specific primer, 5'-CATTTGTGTAGCCTCATCTAGG-3' (SEQ ID NO 37) and 5'-GTCCGGAGAGCAGCCCTTGGTAG-3' (SEQ ID NO 38) were used for 5'RACE, and

5'-TCATCGTATAATTCTTCAGCCTC-3' (SEQ ID NO 39) was used for 3'-RACE.

With 5'RACE, PCR was conducted twice and DNA was obtained. After 8.6 µl of sterilized water, 2 µl of 10x LA PCR buffer solution (manufactured by Takara Shuzo), 2 µl of 25 mM MgCl<sub>2</sub>, 3.2 µl of 2.5 mM dNTP mix, 1 µl of 10 µM primer solution of SEQ ID NO 9907F, 1 µl of 10 µM adaptor primer solution (Marathon RACE system 5'RACE 3'RACE kit, manufactured by Clontech) and 0.2 µl of 5 unit/µl TaKaRa LA Taq (manufactured by Takara Shuzo) were added into 2 µl 250X diluted ogura rapeseed cDNA solution and they were mixed, DNA was amplified by repeating 5 times with a cycle at 98 degrees for 5 seconds and at 72 degrees C for 3 minutes , and 5 times with a cycle at 98 degrees C for 5 seconds and at 70 degrees C for 3 minutes 3, and 25 times with a cycle at 98 degrees C for 5 seconds and at 68 degrees C for 3 minutes. The obtained DNA solution was diluted by 100x, and after 8.6 µl of sterilized water, 2 µl of 10x LA PCR buffer solution (manufactured by Takara Shuzo), 2 µl of 25 mM MgCl<sub>2</sub>, 3.2 µl of 2.5 mM dNTP mix, 1 µl of 10 µM primer solution of SEQ ID NO 5'ogu-1, 1 µl of 10 µM adapter primer solution (Marathon RACE system 5'RACE 3'RACE kit, manufactured by Clontech) and 0.2 µl of 5 unit/µl TaKaRa LA Taq (manufactured by Takara Shuzo) were added to 2 µl out of the obtained diluted DNA solution, DNA was amplified by repeating 5 times with a cycle at 98 degrees C for 5 seconds and at 72 degrees C for 3 minutes, and 5 times with a cycle at 98 degrees for 5 seconds and at 70 degrees C for 3 minutes 3 *[sic.]*, and 25 times with a cycle at 98 degrees C for 5 seconds and at 68 degrees C for 3 minutes. As a thermal cyclers, UNOII (manufactured by Biometra) was used.

With 3'RACE, 8.6 µl of sterilized water, 2 µl of 10x LA PCR buffer solution (manufactured by Takara Shuzo), 2 µl of 25 mM MgCl<sub>2</sub>, 3.2 µl of 2.5 mM dNTP mix, 1 µl of 10 µM primer solution with SEQ ID NO 3'ogu-1, 1 µl of 10 µM adapter primer solution (Marathon RACE system 5'RACE 3'RACE kit, manufactured by Clontech), 0.2 µl of 5 unit/µl TaKaRa LA Taq

(manufactured by Takara Shuzo) were added into 2  $\mu$ l of ogura rapeseed cDNA solution by 50x dilution and they were mixed, DNA was amplified by repeating 35 times with a cycle at 98 degrees C for 5 seconds, at 63 degrees C for 30 seconds and at 72 degrees C for 2 minutes.

Addition of 1  $\mu$ l of pGEM-Teasy vector (manufactured by Promega), 5  $\mu$ l of 2x ligation buffer solution (manufactured by Promega) and 1  $\mu$ l of T4 DNA ligase (manufactured by Promega) into 3  $\mu$ l of the amplified product was conducted, and the mixture was left to stand at room temperature for one hour and this was combined to the vector. This vector was transformed to *Escherichia coli* DH5 $\alpha$  (manufactured by Gibco BRL) and clones were obtained. After base sequences of the obtained clones were determined, 5'-RACE sequences and 3'-RACE sequences corresponding to the four types of cDNA partial sequences were obtained, and each sequence was combined and whole-length *[sic.]* cDNA sequence was obtained. The sequence with the highest homology with the cDNA sequence of this Yuanhong radish (*Raphanus sativus* cv. Yuanhong) was regarded as the fertility restoring gene of kosena rapeseed (SEQ ID NO 18). In addition, amino acid sequences were obtained from the cDNA sequences (SEQ ID NO 19).

#### Embodiment 15: Analysis of fertility restoring gene

Regarding each of the fertility restoring genes obtained with the above-mentioned methods, the homology of the cDNA sequences & the amino acid sequences and the motif were analyzed using the gene analysis software, "Mac Vector 6.5" (Oxford Molecular Ltd.) was used, and Protein families database of alignments and HMNs situated at Sanger Institute, UK, respectively.

Rf genes derived from Yuanhong radish (*Raphanus sativus* cv. Yuanhong), kosena rapeseed and ogura rapeseed all 16 PPR motifs, and this PPR motif group was divided into three blocks, having 5, 7 and 4 motifs, from the amino terminal side, and the fourth amino acid existing in the second

PPR motif from the amino terminal side was asparagine. Further, the partial fragments derived from *Raphanus raphanistrum* were analyzed by applying the above-mentioned sequences to portions equivalent to the above-mentioned sequences (the 94th to 264th in SEQ ID NO 26), and they have portions equivalent to the first to sixth PPR motifs from the amino terminal, and the PPR motif group and the fourth amino acid existing in the second PPR motif from the amino terminal side had the above-mentioned characteristic.

In addition, among the protein of the present invention (SEQ ID NO 26) obtained as a result of analysis of the fertility restoring gene partial sequences of three types of fertility restoring genes obtained from Yuanhong radish (*Raphanus sativus* cv. Yuanhong), kosena rapeseed and ogura rapeseed and gene coding them (SEQ ID NO 22), the homology between Yuanhong radish (*Raphanus sativus* cv. Yuanhong) and kosena rapeseed was very high, and since the values were 99.6 % in the amino acid sequence (a difference was merely confirmed in three sections of amino acid) and 99.7 % in the base sequences, the protein having the amino acid sequence described in SEQ ID NO 29 and a gene coding this protein (SEQ ID NO 25) are particularly preferable.

In the meantime, comparing between ogura rapeseed and Yuanhong radish (*Raphanus sativus* cv. Yuanhong) and between ogura rapeseed and kosena rapeseed, high homology was confirmed in both comparisons, and their homology values were 92.0 % and 91.6 % in the amino acid sequences of comparison between the ogura rapeseed and Yuanhong radish (*Raphanus sativus* cv. Yuanhong) and between ogura rapeseed and kosena rapeseed, respectively, and the value in the base sequence was approximately 95 %, and the values were lower than those between Yuanhong radish (*Raphanus sativus* cv. Yuanhong) and kosena rapeseed.

Industrial Usability

According to the present invention, the Rf genes, particularly the Rf gene derived from radish was isolated and its structure was identified. In addition, according to the present invention, it has become possible to provide a means to establish a means to establish the rapeseed restoration line by utilized isolated Rf gene.

## CLAIMS

1. Proteins participating in restoration from cytoplasmic male sterility to fertility, wherein the proteins contain 14 or more of pentatricopeptide repeat (hereafter, abbreviated as PPR) motifs; the PPR motif group is divided into three or more blocks; each block has at least two or more PPR motifs, respectively; and the block at a carboxy terminal (C-terminal) has four PPR motifs.
2. The proteins according to claim 1, where the number of PPR motif is 14 to 16.
3. The proteins according to claim 1 or 2, wherein the PPR motif group is divided into three blocks, and the blocks have five, seven and four PPR motifs in respective order from the amino terminal (N-terminal) side, respectively.
4. The proteins according to claims 1 to 3, wherein the fourth amino acid existing in the second PPR motif existing from the amino terminal (N-terminal) side is other than serine, threonine and cysteine.
5. The proteins according to claim 4, wherein the fourth amino acid existing in the second PPR motif from the amino terminal (N-terminal) is any of asparagine, glutamine, asparagine acid, glutamic acid and histidine.
6. The proteins according to any of claims 1 to 5 further containing a signal peptide sequence for migrating to mitochondria at the amino terminal, or a sequence comprising -LysAspGluLeu- at the carboxy terminal.

7. Proteins participating in restoration from cytoplasmic male sterility to fertility causing an occurrence of a gel shift to the transcription after making contact with a transcription of a cytoplasmic male sterile gene.
8. Proteins participating in restoration from cytoplasmic male sterility to fertility having amino acid sequences described in SEQ ID NO 26.
9. Proteins participating in restoration from cytoplasmic male sterility to fertility having amino acid sequences described in SEQ ID NO 27.
10. Proteins participating in restoration from cytoplasmic male sterility to fertility having amino acid sequences described in SEQ ID NO 28.
11. Proteins participating in restoration from cytoplasmic male sterility to fertility having amino acid sequences described in SEQ ID NO 29.
12. Any of the following proteins:
  - (1) a protein having the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 3, the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 17, or the 82nd to 690th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 19; or
  - (2) a protein having an amino-acid sequence where one to multiple amino acids are deleted, added and/or substituted in the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 3, the 80th to 687th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 17 or the 82nd to 690th amino-acid sequences among the amino-acid sequences described in SEQ ID NO 19 and participating in restoration from cytoplasmic male sterility to fertility.

13. Any of the following proteins:

(1) a protein having amino-acid sequences described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19; or

(2) a protein having an amino-acid sequence where one to multiple amino acids are deleted, added and/or substituted described in SEQ ID NO 3, SEQ ID NO 17 or SEQ ID NO 19 and participating in restoration from cytoplasmic male sterility to fertility.

14. The proteins according to claim 1 to 9, wherein the cytoplasmic male sterile individual has a cytoplasmic male sterile gene of kosena radish (*Raphanus sativus* cv. Kosena) and/or Ogura radish (*Raphanus sativus* L.) or homologue thereof.

15. DNA coding for the proteins according to any of claims 1 to 10.

16. DNA having base sequences described in SEQ ID NO 22.

17. DNA having base sequences described in SEQ ID NO 23.

18. DNA having base sequences described in SEQ ID NO 24.

19. DNA having base sequences described in SEQ ID NO 25.

20. Any of the following DNA:

(1) DNA having base sequences described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18, or ; or

(2) DNA having base sequences where one to multiple bases are deleted, added and/or substituted in the base sequences described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18 and participating in restoration from cytoplasmic male sterility to fertility; or

(3) DNA hybridizing with base sequences described in SEQ ID NO 2, SEQ ID NO 16 or SEQ ID NO 18 under stringent conditions and participating in restoration from cytoplasmic male sterility to fertility.

21. Any of the following DNA:

(1) DNA having the 3,754th to 8,553th base sequences among the base sequences described in SEQ ID NO 1 or the 812th to 3002nd base sequences among the base sequences described in SEQ ID NO 15; or

(2) DNA having the 3,754th to 8,553th base sequences where one to multiple bases are deleted, added and/or substituted among the base sequences described in SEQ ID NO 1 or the 812th to 3002nd base sequences where one to multiple bases are deleted, added and/or substituted among the base sequences described in SEQ ID NO 15 and participating in restoration from cytoplasmic male sterility to fertility; or

(3) DNA hybridizing with DNA having the 3,754th to 8,553th base sequences among the base sequences described in SEQ ID NO 1 or the 812th to 3002nd base sequences among the base sequences described in SEQ ID NO 15 under stringent conditions and participating in restoration from cytoplasmic male sterility to fertility.

22. Any of the following DNA:

(1) DNA having base sequences described in SEQ ID NO 1 or SEQ ID NO 15; or

(2) DNA having base sequences where one to multiple bases are deleted, added and/or substituted among the base sequences described in SEQ ID NO 1 or SEQ ID NO 15, and participating in restoration from cytoplasmic male sterility to fertility; or

(3) DNA hybridizing with DNA having base sequences described in SEQ ID NO 1 or SEQ ID NO 15 under stringent conditions and participating in restoration from cytoplasmic male sterility to fertility.

23. The DNA according to any one of claims 15 to 22, wherein the cytoplasmic male sterile individual has a cytoplasmic male sterile gene of kosenia radish (*Raphanus sativus* cv. Kosenia) and/or Ogura radish (*Raphanus sativus* L.) or has homologue thereof.
24. A vector containing the DNA according any of claims 15 to 23.
25. A transformant having the DNA according to any of claims 15 to 23 or the vector according to claim 24.
26. The transformant according to claim 25 being a transformed plant.
27. A method for restoring from cytoplasmic male sterility to fertility using the DNA according to any of claims 15 to 23.
28. A transformant enabling the control of expression of a cytoplasmic male sterile gene by partially or entirely introducing the DNA described in any of claims 15 to 23 to a cell having a cytoplasmic male sterile gene and having the DNA according to any of claims 15 to 23.
29. A maintenance method for a cytoplasmic male sterile line by utilizing the transformant according to claim 28.
30. A method for detecting a gene participating in restoration from cytoplasmic male sterility by using 15 mer to 50 mer of oligonucleotide primer arbitrarily set from the DNA according to any of claims 15 to 23 or at least 15 mer or more of probe comprising all or part of the DNA according to claims 15 to 23, and by confirming that a quantity of base sequences to be

amplified by the primer or a quantity of base sequences detected by the probe is one gene or more per genome in a target biological sample.

31. A promoter DNA having the 3754th to 5091st base sequences among the base sequences described in SEQ ID NO 1 or the 1st to 811th base sequences among the base sequences described in SEQ ID NO 15.

FIG. 1

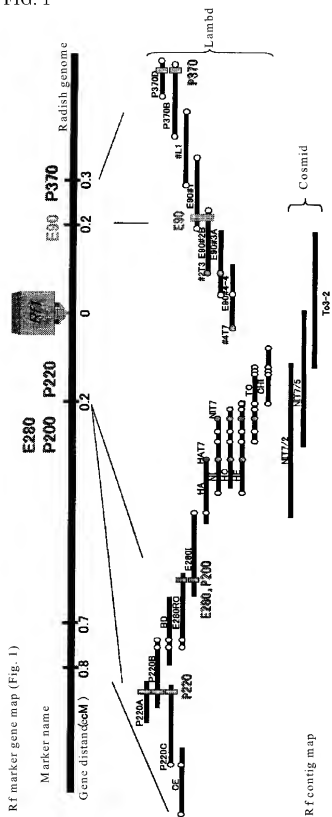


FIG. 2

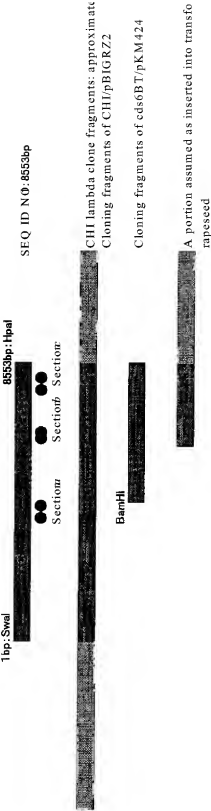


FIG. 3

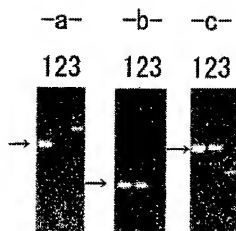


FIG. 4

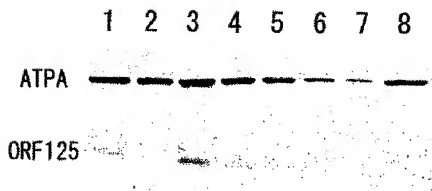


FIG. 5

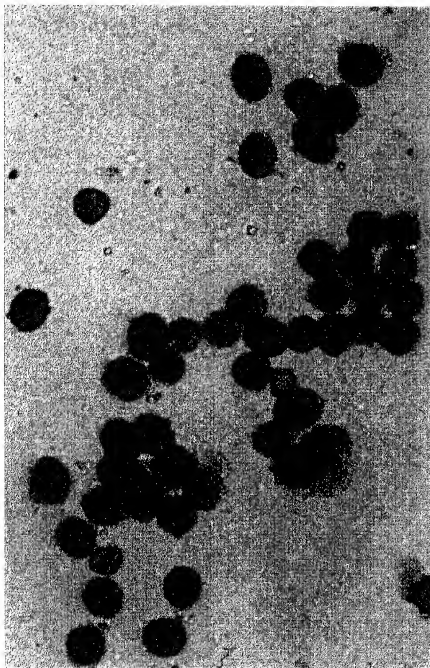


FIG. 6

```

pSTV125-5' #LA12.nuc      1:GGATCCCAATTTCATTCTGCATCACTTCCTCTCGTTATGGACCTGGCAAGTTTTC 60
pSTV125-5' #LA6.nuc      1:GGATCCCAATTTCATTCTGCATCACTTCCTCTCGTTATGGACCTGG-CAAGTTTTC 59
      ***** * *****
pSTV125-5' #LA12.nuc      61:AAACGGCCGAAACGGGAAGTGACAATACCGCTTTCTTGAGCATATAAATCATGATTAC- 119
pSTV125-5' #LA6.nuc      60:AAACGGCCGAAACGGGAAGTGACAATACCGCTTTCTTGAGCATATAAATCATGATTAC 119
      ***** * *****
pSTV125-5' #LA12.nuc      120:CTTTTTCGAAAAATGTCCACTTTTGTGCATATCCTCACTTCTCTACTGAAGTAAAGT 179
pSTV125-5' #LA6.nuc      120:CTTTT-TCGAAAAATGTCCACTTTTGTGCATATCTCACTCTCTACTGAAATAA-A-GT 176
      ***** * *****
pSTV125-5' #LA12.nuc      180:TAGTGAATTC 189
pSTV125-5' #LA6.nuc      177:TAGTGAATTC 186
      *****

```

## SEQUENCE LISTING

&lt;110&gt; Mitsubishi Chemical Corporation

&lt;120&gt; A protein which is involved in recovery of cytoplasm male fertility from sterility and a gene encoding the protein

&lt;130&gt; A21220A

&lt;160&gt; 39

&lt;210&gt; 1

&lt;211&gt; 8553

&lt;212&gt; DNA

<213> *Raphanus sativus*

&lt;400&gt; 1

```

atttaaattt tataactaat atgtatttaa actctccaat gcaataaggg atataaacaa 60
aaggattcca tagatgttat gtattcgtac accgatgtat tcgtatacct taaatatatg 120
tataacttatg tatacatata cttgtgtatt cgtacacett aagtattcga tgggttatgt 180
tggtattcgt atattttatg tatttgcaca ccttatgtat acttatgtat atgtacacct 240
tatgtatttg tacatcttaa gtattagatg agttatgttg atattcgtac accttatgta 300
ttcgtacaca ttctgtatac cttaggtaatt cgtacacett aggtatttgt acacetaagg 360
tattcgtaca ccttatgtat acttatgtat acgtacacct tatatatcgc aacaccttag 420
atattcgtac atcttatgta tacgtatact tattiictga gttatagtga attagattgt 480
attaaacgtt agacataggg ttccggattt atccaagggt tccagattgt ttcagattct 540
ggatttaccg aatggttctg gatltaccga aggtttccgg atttaggatt caaggttttag 600
agtttaggat tttaggttta gtgttttgtt gatgattttt aatatttaag ataaatgtag 660
acaaatttgt tcttcctacc attttgacaa aaaaigaaag atctatgtag gtttccaagt 720
ttattaaatt taccagatt tatgaaaatt atccataaat ttatataatt ttatgaataa 780
tttatcattt atttgggtaa atttcataaa tatgaaagtt ttttttatgg gtcaaatgt 840
ataattttat cggattcigg atttacccaa gggttccgga ttacccaag gattccagat 900
ttaggattca tggtttagag tttaggagtt tatgtttagt gttttgtga tgattttaaa 960

```

tatttaugat aagaagttaa tgcgagagaa ttigtcaaa ctacaggtga gtcttaactt 1020  
cttaagacat aaaaactact agataactga catggaggca ccaaattatc ctatatattt 1080  
tggacttaat ctigtgtgat cctagagta aaccttaagg ttcaccaacc aatagaaatc 1140  
ctctatttca cagttgatat ctittaaaaa agtaaacaaa atattgtcga gttatattac 1200  
atttttaaaa taaaaatatt aaaaaataaa aataataata tatgcacaaa aaaagatttt 1260  
ttaaaaagat ttaattttcg tcaacaaaac actaaacict aaactotaaa tcttaaaccc 1320  
ttggataaat actaaacoot aaattaaaaa cattaaacca taatagtatt ttaagatttt 1380  
aatgttttag tgttttagtt ttttgattta gaatttagga ttatccaagt gtttatgatt 1440  
tatccaaggg tttagggttt agaatttagg gtttagggtt tagagttaa aattatccaa 1500  
gggtctaggg tataccaagg ggtttagggt ttaggattta gggtttaggg tttagaattt 1560  
agggttttag gtttagagtt taaaattatc caagggttta gggttatccc aagggttttag 1620  
ggtttaggat ttagggttta aggttttagt ttttttagc ataltaaaaa tagttttcaa 1680  
aaattcattt ttgttaacgg ctattttttt ttttttatat tttattttt ttaaaaacat 1740  
aatataactt gacaatatit tcttttcttt tttaaaaaa tatlaattat gaataacttg 1800  
attcctattg ttgggttgaa cctaaatgtt cactctaggg gtgaacctaa ggataactct 1860  
attttttggg gtgaatatgc actatagcgg atatcttttl caatagatta taagcacggc 1920  
tctacctatg actaatcaag aacttgggat gattggaaat ctgcaggttg tactcaatat 1980  
gggattatat tggttctaac aagtagatat gatccttgaa aattaaagtt attagatcac 2040  
ttcatctga aagggttagg gtttgtcatt ttattaacaa atttgicatt tcaataacaa 2100  
ttttgtcat tttataacaa tgaanaattat aacgaatgca ctttgtgcc agatcccaat 2160  
ttgtcatttt atttttggga aaaaaatgta gcatttcgtg agtgtttcta tttttggcaa 2220  
aaacaaaagg tgtgagatca attttgacca aaaaaaatg taagattcac gtaggtttcc 2280  
aaatttatta aatttaccca actatattaa aattaaatgt agacaaattt gttttcctgc 2340  
cattttggca aaaaatgaag gatctatgaa ggtttccaag tttattaaat ttaactagat 2400  
ttatgatat tatocataaa ttacataaat tttatgaatt atcattttatt tgggtagatt 2460  
tcataaatat gaaagtttct tttatgagtc aaaaagtata atttatggg taactttcat 2520  
aaattttaga atttacatcg attttatatt aattcgtata gatttatgtt gactttatat 2580

atgaaaaaat atgtattata tlaaagtag ttgtctatat atgattttta aatattaaat 2640  
atgatccaaa agittantga ataaageatg ttlatgggat ttacaaaagt tagttgttaa 2700  
aagttagtgg gaaaaaaatt attttttata ggc aaaagtggtg attttgggtc ccacgaattt 2760  
acttttccaa ctgtccaaagt ttaataggca aaaaggttaa aaatgtcata aattttattct 2820  
ctctctacta ggttgcccaa ttgctaata taaccttgag gtggcctatt ttctaatc 2880  
aaacttaaaa gtggcccttt cccctaattg acccataaaa gaatgaasga caattttctt 2940  
ttccaaatta caatccctag ataattttat ttgtaggta cattccatcg gttatgatta 3000  
cagaatagct acgtctctct attgattctt attgcgcgt tggtagcgtt ttccatggaa 3060  
tcaagtatgt ttttatctoc tactactaac aacatatcca tagattttgt ttatcaottg 3120  
ttctgtgttc ctgatcatat acttgactca gtttctgtga ttctatcaag tttttgagan 3180  
cagaagaago aaaaaagaaa acgagcagag ctgctcttac aatgttttaa cgttgagta 3240  
taaatattat tacataaaag ttttlaaaa atagatttaa tcaaccaatt taatatatta 3300  
ttttatattt agttcatttt ttttgacat cttttatatt tagtttagaa caccctattt 3360  
tgagtacac atagattata atgataaatt tataaaatag cataattttt tttttcatt 3420  
gttttatgat aaattctaa ataacaaata ttataatatt attatattac taattgonaa 3480  
aattanttaa tacattattt tataataaat atttaaaacg ttgggtgga tttgttaga 3540  
ttttttcaa caaattttgt tatagctaaa ataaaattca aatgtattgt taaaattgat 3600  
ttttttttt tttgattatt aagatttaat ataaataaac atatatgca tattaaatat 3660  
ttaactaagt ggtcctaate ttgaactag ggggtggcgt tccggtaact attcgggtt 3720  
cggttcagat ctattcggat ttoggatttt tgggtcaaa gattttagcc caattcgggt 3780  
atttctaant tccggttcgg gtccggttcg gatcctgtcg gatccggttc gggttcggt 3840  
aaccogitta aattattttc aaaattttta aatttcatta tatattttta acttttcgaa 3900  
atttgtaac aaaataatat attaatata aatttcaata atattgtcg aagtacaaa 3960  
acttaacatg taatttgatt tgatttggat atttgatag aaatcaatc atattttata 4020  
tatttttgtt gttttgagta tgatttaact atttatacat gtacttttta atgtttttat 4080  
atattttcta gtattttgaa caatttaaaa gtattatata ttttttagat gctttttat 4140  
atatattcaa tctaaaaata gttaaaata tatgtatatt aatctatttc gatacattc 4200

ggataacaaa aatatatttg ttoggetcgg gticggtttt ggittctttaa ataccacaaa 4260  
tttaaaccta ttoggetatt caattaatit eggttoggat ttgtattacc ttttgcagat 4320  
eggattcggg ttoggttctt ggattcagtt tttttgtcca gccctactct gaacagtaga 4380  
laaaaaantag aaccttaaat taataggtaa gattttgggt aggtctttct aattagtatg 4440  
gagattctcg attccttctc attgcagtgt gglatgtcca actcattgtt tatgtacata 4500  
tccaatitag tttagagtaa aattgttagt tacttaagag ttgaatgaaa tagggagatg 4560  
tattgatggc caaggttctc ccgaagtaaa taactttgtt tatattitaa gttagcttat 4620  
aacatcaata aaaaatgtcat taactgttc aataaaaatg tcattactcg gtctctctaa 4680  
tataattatt taacacacct ggctgttgat aaatttttat gatcgtttaa taattttaga 4740  
agtggatagt ctgtaaatgg tctttgattg gtctcttga tttttaaag tggactaac 4800  
aagaaggctt agtaataaat actgaaccgg aactctactg gtttcaatag ctcggtttat 4860  
caatttctct cgcctctggg tttagtgaat catgtggccc tgtgggttaa aacaaggaac 4920  
tcaatcaatc aactggtagc aactctgaac cggaaattgt ataattcaaa ctgaaccggt 4980  
tcttgtaaaa caaatggaa cgtttgtac ttatctctc gtttatttc tcagtcacga 5040  
gtttttttta gagatcgacg aagaacaaaa tttaggcga acaaaaaataa aatgttggt 5100  
agggtttgtg gattcaagtg ttctttctct cctgtgagt ctgoggttag attgttctgt 5160  
acgagatcga ttcttgatac ttggccaag gcnagcggag agagttcga agcaggttt 5220  
ggagggagaga gtttgaagc gcaaatggg ttatgaana tcaaaagttt agaggatcg 5280  
aattgtttgt tcagtgacat gttcagatct cgtctttac ctctgttgt tgatttctgt 5340  
aaattgatgg gtgtggtgtt gagaatgaa cgcocggatc ttgtgatttc tctctatcag 5400  
aagatggaaa ggaacagat tgatgtgat atatacagct tcaatattct gataaatgt 5460  
ttctgcagct gctctaaat cccctttgct ttgtctaac ttggttaagt caccagctt 5520  
ggaetccaac ctgatgtgt tacttcaac accctgtctc atggattatg tgtggaagt 5580  
agggtttctg aagccttga ttttttcat caaatgttg aaacgacat taggccaat 5640  
gtctaacct tcaccacttt gatgaacggt ctgtgcgcg agsgtagaat tctgaagcc 5700  
gtagctctgc ttgatcgat gatggaagt ggtctccag ctaccagat tacttatgga 5760  
acaatcgtag atgggatgtg taagaaggga gatacttgt ctgactgaa tctgtgagg 5820

aagatggagg aggtgagcca catcataccc aatgttgtaa tctatagtgc aatcattgat 5880  
agcctttgta aagacggacg tcatagcgat gcacaaaac tttcaactga antgcaagag 5940  
aaaggaatct ttcccgattt atttacctac aacagtatga tagttggttt ttgtagctct 6000  
ggtagatgga ggcacgggga gcagttgttg caagaaatgt tagaaaggaa gatcagccct 6060  
gatgttgtaa cttataatgc ttgatcaat gcatttgtca aggaaggcaa gttctttgag 6120  
gcigaaagat tatacgatga gatgcttoca aggggtataa tcctaatac aatcacalat 6180  
agttcaatga togtatgatt ttgcaaacag aatcgtcttg atcgtctgta gcacatgttt 6240  
tattttgatg ctaccaaggg ctgctctccc aacctaatca cttcaatac tctcatagac 6300  
ggatatgttg gggctaagag gatagatgat ggaalggaa ttctocatga gatgactgaa 6360  
acagatttag ttgctgacac aactacttac aacactctta ttcacgggtt ctatctggtg 6420  
ggcgatctta atgctgctct agacctitta caagagatga tctctatggt ttgtgccct 6480  
gatatogtta ctgtgacac ttgtctggat ggtctctgag ataattggaa actaaagat 6540  
gcaatggaa tggttaaggt tatgcagaag agtaagaagg atcttgatgc tagtcacccc 6600  
ttcaatgggt tggaaacctg tgttcaaac tacaatatat tgatcagcgg cttgatcaat 6660  
gaagggaagt ttttagagcg cgggaatta tacgaggaga tgcccacag ggttatagtc 6720  
ccagatacta tcacctatag ctcaatgate gatggattat gcaagcagag ccgcctagat 6780  
gaggotacac aatgtttga ttogatgggt agcaaggagct tctctccaa cgtatgacc 6840  
tttactacac tcattaatgg ctactgtaag gcagggaagg ttgatgatg gctggagcct 6900  
ttctgcgaga tgggtcgaa aggatagtt gctaacgca ttacttacat cactttgatt 6960  
tgtgttttc gtaagtggg taatatlaat ggggccttag acattttcca gtagatgatt 7020  
tcaagtgttg tctatcctga taconttacc atccgcaata tgctgacgg tttatggagt 7080  
aaagaggaa ctaaaaggcg agtggcaatg cttagaanaa tgcagatgag tatggtatgt 7140  
aagtttctgt tcagtctatg tattttttat ataaacaga atgtatacat tcttttgtgt 7200  
gtagcttcag attgatgata cagcttctgg aatlaacct tggtttggtt ttgcattgta 7260  
ggtatctatca ttgggggagt gaatgatcaa agattttctt ctgtttgcgc agcagagctt 7320  
caatgtcatt ttgtttctgc tgctgcattg atacctact aatgittgat caaatcgttg 7380  
aatagagtga tcatagtga aaattgtgtg gttagtaagt tattttgctg ctattcta 7440

gaccgcccitt tatgcgtctc ttgtctgggc ttaataaatt tgaccatttc caattaaatt 7500  
 ccatacaact gtttcacgca agattattgg tctgaactaa agaggcacac ctccagag 7560  
 atttcagggtg ttaaaagatg tttagggtgc tgcgcgttct gtagctgtca ccattggtat 7620  
 cgtccagctc ggtcttcacg agagctgata gctgtgatgc catcttcctc cttctcttca 7680  
 tatgtgctct gtctcgctt gtctgctccc atgtgggttc aggagagat catgttcttt 7740  
 taatcttggg ggaatgttg ttgtcgctta tgcctctctg gttcgctctt tgacttgctt 7800  
 agcttccttc ttatctccca aattgctatg aaatcaattt accataagta gaataaacct 7860  
 gcagattcat tctattattg ctttaagctt tgtaataca caaagaacac agagacgaga 7920  
 aatcacaaat ctataagctt ctcttttttc ttcttgata gtaaacccgg ttagagagta 7980  
 gagattgac atatgaacta aaatcgata ctaaaacggt ttggtccga ctataaac 8040  
 ggaacccac cgttttgcat ctctctctca aacatcac aatgtccaag atgaagagt 8100  
 atttgtgttg tcatctctct ggtgaggag atcaaatgt tatattctaa ttgttticag 8160  
 tgcttggtct aactttttta agagattact ccagtggtt ggaacaaaga aagagtcaac 8220  
 attgcattgt gtaaggtgac gaaactgag tttaaagtaag tgagaacaat acttcaatgc 8280  
 tttctttgtg acaacctgtg taatcatcgc atttgaatat atatgtatat gatgcttatg 8340  
 atgaagctat gagaatagcg aaataggctc tgtgttattt ccttgcgatt ctagattctg 8400  
 atttgttttt cttcttcaat atttagatta ggtggctctg cttatctgt tttagtatta 8460  
 gactcgagat ttgggggatg aatcatcccg gatgatatat acaattgtgt attttatgaa 8520  
 ttctagtttt tagtgataa tgaacacggt aac 8553

<210> 2

<211> 2064

<212> DNA

<213> *Raphanus sativus*

<400> 2

atg ttg gct agg gtt tgt gga ttc aag tgt tct tet tet cct gct gag 48

Met Leu Ala Arg Val Cys Gly Phe Lys Cys Ser Ser Ser Pro Ala Glu

1

5

10

15

```

tct gcg gct aga ttg ttc tgt acg aga tgg att cgt gat act ctg gcc 96
Ser Ala Ala Arg Leu Phe Cys Thr Arg Ser Ile Arg Asp Thr Leu Ala
      20              25              30
aag gca agc gga gag agt tgc gaa gca ggt ttt gga gga gag agt ttg 144
Lys Ala Ser Gly Glu Ser Cys Glu Ala Gly Phe Gly Gly Glu Ser Leu
      35              40              45
aag ctg caa agt ggg ttt cat gaa atc aaa ggt tta gag gat gcg att 192
Lys Leu Gln Ser Gly Phe His Glu Ile Lys Gly Leu Glu Asp Ala Ile
      50              55              60
gat ttg ttc agt gac atg ctt cga tct cgt cct tta cct tct gtg gtt 240
Asp Leu Phe Ser Asp Met Leu Arg Ser Arg Pro Leu Pro Ser Val Val
      65              70              75              80
gat ttc tgt aaa ttg atg ggt gtg gtg gtg aga atg gaa cgc ccg gat 288
Asp Phe Cys Lys Leu Met Gly Val Val Val Arg Met Glu Arg Pro Asp
      85              90              95
ctt gtg att tct ctc tat cag aag atg gaa agg aaa cag att cga tgt 336
Leu Val Ile Ser Leu Tyr Gln Lys Met Glu Arg Lys Gln Ile Arg Cys
      100             105             110
gat ata tac agc ttc aat att ctg ata aaa tgt ttc tgc agc tgc tct 384
Asp Ile Tyr Ser Phe Asn Ile Leu Ile Lys Cys Phe Cys Ser Cys Ser
      115             120             125
aag ctc ccc ttt gct ttg tct aca ttt ggt aag atc acc aag ctt gga 432
Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys Ile Thr Lys Leu Gly
      130             135             140
ctc caa cct gat gtt gtt acc ttc acc acc ctg ctc cat gga tta tgt 480
Leu His Pro Asp Val Val Thr Phe Thr Thr Leu Leu His Gly Leu Cys
      145             150             155             160

```

```

gtg gaa gat agg gtt tct gaa gcc ttg gat ttt ttt cat caa atg ttt 528
Val Glu Asp Arg Val Ser Glu Ala Leu Asp Phe Phe His Gln Met Phe
      165              170              175

gaa acg aca tgt agg ccc aat gtc gta acc ttc acc act ttg atg aac 576
Glu Thr Thr Cys Arg Pro Asn Val Val Thr Phe Thr Thr Leu Met Asn
      180              185              190

gat ctt tgc cgc gag ggt aga att gtc gaa gcc gta get ctg ctt gat 624
Gly Leu Cys Arg Glu Gly Arg Ile Val Glu Ala Val Ala Leu Leu Asp
      195              200              205

egg atg atg gaa gat ggt ctc cag cct acc cag att act tat gga aca 672
Arg Met Met Glu Asp Gly Leu Gln Pro Thr Gln Ile Thr Tyr Gly Thr
      210              215              220

atc gta gat ggg atg tgt aag aag gga gat act gtg tct gca ctg aat 720
Ile Val Asp Gly Met Cys Lys Lys Gly Asp Thr Val Ser Ala Leu Asn
      225              230              235              240

ctg ctg agg aag atg gag gag gtg agc cac atc ata ccc aat gtt gta 768
Leu Leu Arg Lys Met Glu Glu Val Ser His Ile Ile Pro Asn Val Val
      245              250              255

atc tat agt gca atc att gat agc ctt tgt aaa gac gga cgt cat agc 816
Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg His Ser
      260              265              270

gat gca caa aat ctt ttc act gaa atg caa gag aaa gga atc ttt ccc 864
Asp Ala Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile Phe Pro
      275              280              285

gat tta ttt acc tac aac agt atg ata gtt ggt ttt tgt agc tct ggt 912
Asp Leu Phe Thr Tyr Asn Ser Met Ile Val Gly Phe Cys Ser Ser Gly
      290              295              300

```

aga tgg agc gac ggc gag cag ttg ttg caa gaa atg tta gaa agg aag 960  
 Arg Trp Ser Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu Arg Lys  
 305 310 315 320  
 atc agc cct gat gtt gta act tat aat gct ttg atc aat gca ttt gtc 1008  
 Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala Phe Val  
 325 330 335  
 aag gaa ggc aag ttc ttt gag gct gaa gaa tta tac gat gag atg ctt 1056  
 Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu Met Leu  
 340 345 350  
 cca agg ggt ata atc cct aat aca atc aca tat agt tca atg atc gat 1104  
 Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met Ile Asp  
 355 360 365  
 gga ttt tgc aaa cag aat cgt ctt gat gct gct gag cac atg ttt tat 1152  
 Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met Phe Tyr  
 370 375 380  
 ttg atg gct acc aag ggc tgc tct ccc sac cta atc act ttc aat act 1200  
 Leu Met Ala Thr Lys Gly Cys Ser Pro Asn Leu Ile Thr Phe Asn Thr  
 385 390 395 400  
 ctc ata gac gga tat tgt ggg gct aag agg ata gat gat gga atg gaa 1248  
 Leu Ile Asp Gly Tyr Cys Gly Ala Lys Arg Ile Asp Asp Gly Met Glu  
 405 410 415  
 ctt ctc cat gag atg act gaa aca gga tta gtt gct gac aca act act 1296  
 Leu Leu His Glu Met Thr Glu Thr Gly Leu Val Ala Asp Thr Thr Thr  
 420 425 430  
 tac aac act ctt att cac ggg ttc tat ctg gtg ggc gat ctt aat gct 1344  
 Tyr Asn Thr Leu Ile His Gly Phe Tyr Leu Val Gly Asp Leu Asn Ala  
 435 440 445

gct cta gac ctt tta caa gag atg atc tct agt ggt ttg tgc oot gat 1392  
 Ala Leu Asp Leu Leu Gln Glu Met Ile Ser Ser Gly Leu Cys Pro Asp  
 450 455 460  
 atc gtt act tgt gac act ttg ctg gat ggt ctc tgc gat aat ggg aaa 1440  
 Ile Val Thr Cys Asp Thr Leu Leu Asp Gly Leu Cys Asp Asn Gly Lys  
 465 470 475 480  
 cta aaa gat gca ttg gaa atg ttt aag gtt atg cag aag agt aag aag 1488  
 Leu Lys Asp Ala Leu Glu Met Phe Lys Val Met Gln Lys Ser Lys Lys  
 485 490 495  
 gat ctt gat gct agt cac ccc ttc aat ggt gtg gaa cct gat gtt caa 1536  
 Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp Val Gln  
 500 505 510  
 act tac aat ata ttg atc agc ggc ttg atc aat gaa ggg aag ttt tta 1584  
 Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys Phe Leu  
 515 520 525  
 gag gcc gag gaa tta tac gag gag atg ccc cac agg ggt ata gtc cca 1632  
 Glu Ala Glu Glu Leu Tyr Glu Glu Met Pro His Arg Gly Ile Val Pro  
 530 535 540  
 gat act atc acc tat agc tca atg atc gat gga tta tgc aag cag agc 1680  
 Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys Gln Ser  
 545 550 555 560  
 cgc cta gat gag gct aca caa atg ttt gat tgc atg ggt agc aag agc 1728  
 Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser Lys Ser  
 565 570 575  
 ttc tct cca aac gta gtg acc ttt act aca ctc att aat ggc tac tgt 1776  
 Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Asn Gly Tyr Cys  
 580 585 590

aag gca gga agg gtt gat gat ggg ctg gag ctt ttc tgc gag atg ggt 1824  
 Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu Met Gly

595

600

605

cga aga ggg ata gtt gct aac gca att act tac atc act ttg att tgt 1872  
 Arg Arg Gly Ile Val Ala Asn Ala Ile Thr Tyr Ile Thr Leu Ile Cys

610

615

620

ggt ttt cgt aaa gtg ggt aat att aat ggg gct cta gac att ttc cag 1920  
 Gly Phe Arg Lys Val Gly Asn Ile Asn Gly Ala Leu Asp Ile Phe Gln

625

630

635

640

gag atg att tca agt ggt gtg tat cct gat acc att acc atc cgc aat 1968  
 Glu Met Ile Ser Ser Gly Val Tyr Pro Asp Thr Ile Thr Ile Arg Asn

645

650

655

atg ctg act ggt tta tgg agt aaa gag gaa cta aaa agg gca gtg gca 2016  
 Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Ala Val Ala

660

665

670

atg ctt gag aaa ctg cag atg agt atg gat cta tca ttt ggg gga tga 2064  
 Met Leu Glu Lys Leu Gln Met Ser Met Asp Leu Ser Phe Gly Gly

675

680

685

&lt;210&gt; 3

&lt;211&gt; 687

&lt;212&gt; PRT

&lt;213&gt; Raphanus sativus

&lt;400&gt; 3

Met Leu Ala Arg Val Cys Gly Phe Lys Cys Ser Ser Ser Pro Ala Glu

1

5

10

15

Ser Ala Ala Arg Leu Phe Cys Thr Arg Ser Ile Arg Asp Thr Leu Ala

20

25

30

Lys Ala Ser Gly Glu Ser Cys Glu Ala Gly Phe Gly Gly Glu Ser Leu  
 35 40 45  
 Lys Leu Gln Ser Gly Phe His Glu Ile Lys Gly Leu Glu Asp Ala Ile  
 50 55 60  
 Asp Leu Phe Ser Asp Met Leu Arg Ser Arg Pro Leu Pro Ser Val Val  
 65 70 75 80  
 Asp Phe Cys Lys Leu Met Gly Val Val Val Arg Met Glu Arg Pro Asp  
 85 90 95  
 Leu Val Ile Ser Leu Tyr Gln Lys Met Glu Arg Lys Gln Ile Arg Cys  
 100 105 110  
 Asp Ile Tyr Ser Phe Asn Ile Leu Ile Lys Cys Phe Cys Ser Cys Ser  
 115 120 125  
 Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys Ile Thr Lys Leu Gly  
 130 135 140  
 Leu His Pro Asp Val Val Thr Phe Thr Thr Leu Leu His Gly Leu Cys  
 145 150 155 160  
 Val Glu Asp Arg Val Ser Glu Ala Leu Asp Phe Phe His Gln Met Phe  
 165 170 175  
 Glu Thr Thr Cys Arg Pro Asn Val Val Thr Phe Thr Thr Leu Met Asn  
 180 185 190  
 Gly Leu Cys Arg Glu Gly Arg Ile Val Glu Ala Val Ala Leu Leu Asp  
 195 200 205  
 Arg Met Met Glu Asp Gly Leu Gln Pro Thr Gln Ile Thr Tyr Gly Thr  
 210 215 220  
 Ile Val Asp Gly Met Cys Lys Lys Gly Asp Thr Val Ser Ala Leu Asn  
 225 230 235 240  
 Leu Leu Arg Lys Met Glu Glu Val Ser His Ile Ile Pro Asn Val Val

245	250	255
Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg His Ser		
260	265	270
Asp Ala Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile Phe Pro		
275	280	285
Asp Leu Phe Thr Tyr Asn Ser Met Ile Val Gly Phe Cys Ser Ser Gly		
290	295	300
Arg Trp Ser Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu Arg Lys		
305	310	315
Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala Phe Val		
325	330	335
Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu Met Leu		
340	345	350
Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met Ile Asp		
355	360	365
Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met Phe Tyr		
370	375	380
Leu Met Ala Thr Lys Gly Cys Ser Pro Asn Leu Ile Thr Phe Asn Thr		
385	390	395
Leu Ile Asp Gly Tyr Cys Gly Ala Lys Arg Ile Asp Asp Gly Met Glu		
405	410	415
Leu Leu His Glu Met Thr Glu Thr Gly Leu Val Ala Asp Thr Thr Thr		
420	425	430
Tyr Asn Thr Leu Ile His Gly Phe Tyr Leu Val Gly Asp Leu Asn Ala		
435	440	445
Ala Leu Asp Leu Leu Gln Glu Met Ile Ser Ser Gly Leu Cys Pro Asp		
450	455	460

Ile Val Thr Cys Asp Thr Leu Leu Asp Gly Leu Cys Asp Asn Gly Lys			
465	470	475	480
Leu Lys Asp Ala Leu Glu Met Phe Lys Val Met Gln Lys Ser Lys Lys			
	485	490	495
Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp Val Gln			
	500	505	510
Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys Phe Leu			
	515	520	525
Glu Ala Glu Glu Leu Tyr Glu Glu Met Pro His Arg Gly Ile Val Pro			
	530	535	540
Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys Gln Ser			
545	550	555	560
Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser Lys Ser			
	565	570	575
Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Asn Gly Tyr Cys			
	580	585	590
Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu Met Gly			
	595	600	605
Arg Arg Gly Ile Val Ala Asn Ala Ile Thr Tyr Ile Thr Leu Ile Cys			
	610	615	620
Gly Phe Arg Lys Val Gly Asn Ile Asn Gly Ala Leu Asp Ile Phe Gln			
625	630	635	640
Glu Met Ile Ser Ser Gly Val Tyr Pro Asp Thr Ile Thr Ile Arg Asn			
	645	650	655
Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Ala Val Ala			
	660	665	670
Met Leu Glu Lys Leu Gln Met Ser Met Asp Leu Ser Phe Gly Gly			

675	680	685
<210> 4		
<211> 25		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Description of Artificial Sequence: Synthetic DNA		
<400> 4		
gaagcaaaa agaaaacgeg cagag'	25	
<210> 5		
<211> 25		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Description of Artificial Sequence: Synthetic DNA		
<400> 5		
ccaaaaatcc gaatatcgaa tagac	25	
<210> 6		
<211> 20		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Description of Artificial Sequence: Synthetic DNA		
<400> 6		
ctcggtctctg ggtttagtga	20	
<210> 7		
<211> 20		

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 7

tccacaaacc ctacccaaca 20

<210> 8

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 8

gcttatgctt ctctggttcg cctc 24

<210> 9

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 9

ctcagttttc gtcacettac acaatgc 27

<210> 10

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 10

gattcctttc tcttgcaatt cag 23

<210> 11

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 11

atctcgtctt ttacctcttg tgg 23

<210> 12

<211> 18

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 12

cgggatccgc tcacaatt 18

<210> 13

<211> 100

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 13

ggggatccca atttcattct gcatcaactct cctgtcgtt atcgaccteg caaggttttt 60

gaaacggcgg aaacgggaag tgacaatacc gctttttcttc 100

<210> 14

<211> 100

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 14

ggaaattcaat aactttacat tcagtaggag tgagattatg acaaaaagtg gacaattttt 60

cguaaaagggt aatcatgcat ttatatgctg aagaaaagcg 100

<210> 15

<211> 3306

<212> DNA

<213> *Raphanus sativus*

<400> 15

caattaattt cggitcggat ttggtattac ttttcagat cggaitcggg toggttcttt 60

ggattcagtt tttttgtcca gccctactct gaacagtaga tasaaaatag aaccctaaat 120

taataaggta gatttttggt aggtctttct aattagratg gagattctcg attccttctc 180

attgcagltg ggtatgtcca actcattggt tatgtacata tccaatttag ttttgagica 240

aatgtttagt tacttaagag ttgaatgaaa taggggatga tattgatgac caaggctctc 300

ccaaagttaa taactttggt tatattttta gttagcttat aacatcaata aaaatgtcat 360

taactgttcc aataaaaatg tcatiaactg gttcctctaa tataattatt taacacacct 420

ggctgttgat aaatttttat gatcgtttta taattttaga agtggatagt ctgtaaatgg 480

tctttgatgg gtgcgtctga tttttaaaag tggactaaac aagaaggctt agtaataaat 540

acigaaacgg aactctactg gtttcaatag ctgggtttat caattttctt cggctctcgg 600

tttagtgaat catgtgcgcc tgtgggitta aacaaggaac tcaatcaatc aactgggtgac 660

aaatctgaac cggaatttgt ataattcaaa ctgaaccggg tcttgtaaaa caaatggaac 720

cagtttgtaac tttatctctc gtttatttllc tcagtcacga gtttttttta gagatogacg 780  
 aagaacaaaa tttaggcgaa acaaaaataa aatgttggct agggtttttg gattcaagtg 840  
 ttctttctct cotgctgagt ctgcggctag attgttctgt acgagatoga ttctgataac 900  
 tcigcccaag gcaagcggag agagttcgga agcaggttll ggaggagaga gtttgaagct 960  
 gcaaaagtggg ttatcatgaa tcaaaaggtt agaggatgag attgattttg tcagtgcacat 1020  
 gcttcgalel cgtcccttac cttctgtggt tgaattctgt aaattgaigg gtgtggtggt 1080  
 gagaatggaa cgcoccgatc ttgtgatttc tctctatcag aagatggaaa ggaaacagat 1140  
 tcgagtgtat atatacagct tcaattctct gataaaatgt ttctgcagct gctctaaagct 1200  
 cccctttgct ttgtctaacat ttggtaaagct caccacagct ggactccacc ctgagtgtgt 1260  
 taccttcaac accctgctcc accgatttgg cgtggaagat agggtttctg aagctttgaa 1320  
 ttgtttctat caaatgtttg aaacgacatg taggccaat gtcgtaacct tcaccacttt 1380  
 gatgaacggg ctttgcgcg agggtagaat tgcgaagcc gtagctctgc ttgatcggtat 1440  
 gatggaagat ggtctccagc ctaccagat tacttatgga acaatcgtag atgggatgtg 1500  
 taagaaggga gatactgtgt ctgcactgaa tctgctgagg aagatggagg agtgagcca 1560  
 catcataccc aatgttgtaa tctatagtgc aatcattgat agcctttgta aagacggagc 1620  
 tcatagcgat gcacaaaaac ttttcaactga aatgcagag aaaggaatct ttcccgattt 1680  
 atttacctac aacagtaiga tagttggttt ttgtagctct ggtagatgga gcgacgcgga 1740  
 gcagttgttg caagaatgt tagaaaggaa gatcagccct gatgttgtaa cttataatgc 1800  
 ttgatcaat gcatttgcga aggaaggcaa gtcttttgag gctgaagaat tatacgatga 1860  
 gatgcttcca aggggtataa tccctaatc aatcacatat agttcaatga tcgatggatt 1920  
 ttgcaaacag aatcgtcttg atgtctctga gcaatgttt tatttgatgg ctaccaaggg 1980  
 ctgctctccc aaactaatc ctttcaatc tctcatagac gcatattgtg gggctaaggag 2040  
 gatagatgat ggaatggaac ttctccatga gatgactgaa acagzattag ttgtgcac 2100  
 aactacttao aacactctta ttacagggtt ctatctgggt ggcgacttta atgtgtctct 2160  
 agacctttta caagagatga tctctagttg ttgtgccct gatatcgta cttgtgacac 2220  
 ttgtctggtt ggtctctgag acaatgggaa actaaagat gcattggaa tgtttaaggt 2280  
 tatgcagaag agtaagaagg atcttgatgc tagtcaccc ttcaatgggt tggaaacctga 2340

tgttcaact tacaatatat tcatcagcgg ctgtatcaat gaagggaagt ttttagaggc 2400  
 cgaggnaatta tacgaggaga tgcgccacag ggggtatagtc ccagatacta tcacctatag 2460  
 ctcaatgctc gatggattat gcaagcagag ccgcctagat gaggctacac aaatgtttgn 2520  
 ttogatgggt agcaagagct tctctccaa cgtagtgcac ttactacac tcattaatgg 2580  
 ctactgtaag gcagggaaggg ttgatgatgg gctggagctt ttctgcgaga tgggtcgaag 2640  
 egggatagtt gctaaacgaa ttacttacct cactttgatt tgttggtttc gtaagtggg 2700  
 taatattaat gggactctag acattttcca ggagatgatt tcaagtgggt tgtaacctga 2760  
 taccattacc atccgcata tgcgactgg tttatggagt aaagaggaaac taaaaggggc 2820  
 agtggcaatg cttgagaana tgcgatgag tatggtatgt aagtttctgt tcaagtctatg 2880  
 tattttttat ataacaaga atgtatacat tcttttgtgt gtagcttcag attgatgata 2940  
 caegtctctg aattaacctt tggtttggt ttgcattgta ggaatcataa ttgggggat 3000  
 gnatgatcna agattttctt ctgtttgcgc agcagagctt caatgtcatt ttgtttctgc 3060  
 tgcgtcatgt atacctact aatgtttgat caaatcgtg aatagagtga tcatagtga 3120  
 aaattgtgtg gttagttaag tatittgotg ctattctaag gacagccttt tatgctcta 3180  
 ttgtctgggc ttaataaatt tgaccatttc caattaastt ccatacaatt gtttcacgca 3240  
 agattattgg tctgaactaa agaggcacac ctccagaag atttcagggt ttaaaagatg 3300  
 tttagg 3306

&lt;210&gt; 16

&lt;211&gt; 2064

&lt;212&gt; DNA

<213> *Raphanus sativus*

&lt;400&gt; 16

atg ttg gct agg gtt tgt gga ttc aag tgt tct tct tct cct gct gag 48  
 Met Leu Ala Arg Val Cys Gly Phe Lys Cys Ser Ser Ser Pro Ala Glu  
 1 5 10 15  
 tct gcg gct aga ttg ttc tgt acg aga tgc att cgt gat act ctg gcc 96  
 Ser Ala Ala Arg Leu Phe Cys Thr Arg Ser Ile Arg Asp Thr Leu Ala

20	25	30	
aag gca agc gga gag agt tgc gaa gca ggt ttt gga gga gag agt ttg			144
Lys Ala Ser Gly Glu Ser Cys Glu Ala Gly Phe Gly Gly Glu Ser Leu			
35	40	45	
aag ctg cca agt ggg ttt cat gaa atc aaa ggt tta gag gat gcg att			192
Lys Leu Gln Ser Gly Phe His Glu Ile Lys Gly Leu Glu Asp Ala Ile			
50	55	60	
gat ttg ttc agt gac atg ctt cga tct cgt cct tta cct tct gtg gtt			240
Asp Leu Phe Ser Asp Met Leu Arg Ser Arg Pro Leu Pro Ser Val Val			
65	70	75	80
gat ttc tgt aaa ttg atg ggt gtg gtg gtg aga atg gaa cgc cgc gat			288
Asp Phe Cys Lys Leu Met Gly Val Val Val Arg Met Glu Arg Pro Asp			
85	90	95	
ctt gtg att tct ctc tat cag aag atg gaa agg asa cag att cga tgt			336
Leu Val Ile Ser Leu Tyr Gln Lys Met Glu Arg Lys Gln Ile Arg Cys			
100	105	110	
gat ata tac agc ttc aat att ctg ata aaa tgt ttc tgc agc tgc tct			384
Asp Ile Tyr Ser Phe Asn Ile Leu Ile Lys Cys Phe Cys Ser Cys Ser			
115	120	125	
aag ctc ccc ttt get ttg tet aca ttt ggt aag ctc acc aag ctt gga			432
Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys Leu Thr Lys Leu Gly			
130	135	140	
ctc cac cct gat gtt gtt acc ttc acc acc ctg ctc cac gga tgg tgc			480
Leu His Pro Asp Val Val Thr Phe Thr Thr Leu Leu His Gly Leu Cys			
145	150	155	160
gtg gaa gat agg gtt tct gaa get ttg aat ttg ttt cat caa atg ttt			528
Val Glu Asp Arg Val Ser Glu Ala Leu Asn Leu Phe His Gln Met Phe			

165	170	175	
gaa aag aca tgt agg ccc aat gtc gta acc ttc acc act ttg atg aac 576			
Glu Thr Thr Cys Arg Pro Asn Val Val Thr Phe Thr Thr Leu Met Asn			
180	185	190	
ggt ott tgc cgc gag ggt aga att gtc gaa gcc gta gct ctg ctt gat 624			
Gly Leu Cys Arg Glu Gly Arg Ile Val Glu Ala Val Ala Leu Leu Asp			
195	200	205	
cag atg atg gaa gat ggt ctc cag cct acc cag att act tat gga aca 672			
Arg Met Met Glu Asp Gly Leu Gln Pro Thr Gln Ile Thr Tyr Gly Thr			
210	215	220	
atc gta gat ggg atg tgt aag aag gga gat act gtg tct gca ctg aat 720			
Ile Val Asp Gly Met Cys Lys Lys Gly Asp Thr Val Ser Ala Leu Asn			
225	230	235	240
ctg ctg agg aag atg gag gag gtg agc cac atc ata ccc aat gtt gta 768			
Leu Leu Arg Lys Met Glu Glu Val Ser His Ile Ile Pro Asn Val Val			
245	250	255	
atc tat agt gca atc att gat agc ctt tgt aaa gac gga cgt cat agc 816			
Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg His Ser			
260	265	270	
gat gca caa aat ctt ttc act gaa atg caa gag aaa gga atc ttt ccc 864			
Asp Ala Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile Phe Pro			
275	280	285	
gat tta ttt acc tac aac agt atg ata gtt ggt ttt tgt agc tct ggt 912			
Asp Leu Phe Thr Tyr Asn Ser Met Ile Val Gly Phe Cys Ser Ser Gly			
290	295	300	
aga tag agc gac gcg gag cag ttg ttg caa gaa atg tta gaa agg aag 960			
Arg Trp Ser Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu Arg Lys			

305	310	315	320
atc agc cct gat gtt gta act tat aat gct ttg atc aat gca ttt gtc 1008			
Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala Phe Val			
325	330	335	
aag gaa ggc aag ttc ttt gag gct gaa gaa tta tac gat gag atg ctt 1056			
Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu Met Leu			
340	345	350	
cca agg ggt ata atc cct aat aca atc aca tat agt tca atg atc gat 1104			
Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met Ile Asp			
355	360	365	
gga ttt tgc asa cag aat cgt ctt gat gct gct gag cac atg ttt tat 1152			
Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met Phe Tyr			
370	375	380	
ttg atg gct acc aag ggc tgc tct ccc aac cta atc act ttc aat act 1200			
Leu Met Ala Thr Lys Gly Cys Ser Pro Asn Leu Ile Thr Phe Asn Thr			
385	390	395	400
ctc ata gac gga tat tgt ggg gct aag agg ata gat gat gga atg gaa 1248			
Leu Ile Asp Gly Tyr Cys Gly Ala Lys Arg Ile Asp Asp Gly Met Glu			
405	410	415	
ctt ctc cat gag atg act gaa aca gga tta gtt gct gac aca act act 1296			
Leu Leu His Glu Met Thr Glu Thr Gly Leu Val Ala Asp Thr Thr Thr			
420	425	430	
tac aac act ctt att cac ggg ttc tat ctg gtg ggc gat ctt aat gct 1344			
Tyr Asn Thr Leu Ile His Gly Phe Tyr Leu Val Gly Asp Leu Asn Ala			
435	440	445	
gct cta gac ctt tta caa gag atg atc tct agt ggt ttg tgc cct gat 1392			
Ala Leu Asp Leu Leu Gln Glu Met Ile Ser Ser Gly Leu Cys Pro Asp			

450	455	460	
atc gtt act tgt gac act ttg ctg gat ggt ctc tgc gat aat ggg aaa 1440			
Ile Val Thr Cys Asp Thr Leu Leu Asp Gly Leu Cys Asp Asn Gly Lys			
465	470	475	480
cta aaa gat gca ttg gaa atg ttt aag gtt atg cag aag egt aag aag 1488			
Leu Lys Asp Ala Leu Glu Met Phe Lys Val Met Gln Lys Ser Lys Lys			
	485	490	495
gat ctt gat gct agt cac ccc ttc aat ggt gtg gaa cct gat gtt caa 1536			
Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp Val Gln			
	500	505	510
act tac aat ata ttg atc agc ggc ttg atc aat gaa ggg aag ttt tta 1584			
Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys Phe Leu			
	515	520	525
gag gcc gag gaa tta tac gag gag atg ccc cac agg ggt ata gtc cca 1632			
Glu Ala Glu Glu Leu Tyr Glu Glu Met Pro His Arg Gly Ile Val Pro			
	530	535	540
gat act atc acc tat agc tca atg atc gat gga tta tgc aag cag agc 1680			
Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys Gln Ser			
545	550	555	560
cgc cta gat gag gct aca caa atg ttt gat tcg atg ggt agc aag agc 1728			
Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser Lys Ser			
	565	570	575
ttc tot cca aac gta gtg acc ttt act aca ctc att aat ggc tac tgt 1776			
Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Asn Gly Tyr Cys			
	580	585	590
aag gca gga agg gtt gat gat ggg ctg gag ctt ttc tgc gag atg ggt 1824			
Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu Met Gly			

595                      600                      605  
 cga aga ggg ata gtt gct aac goa att act tac atc act ttg att tgt 1872  
 Arg Arg Gly Ile Val Ala Asn Ala Ile Thr Tyr Ile Thr Leu Ile Cys  
 610                      615                      620  
 ggt ttt cgt aaa gtg ggt aat att aat ggg gct cta gac att ttc cag 1920  
 Gly Phe Arg Lys Val Gly Asn Ile Asn Gly Ala Leu Asp Ile Phe Gln  
 625                      630                      635                      640  
 gag atg att tca agt ggt gtg tat cct gat acc att acc atc cgc aat 1968  
 Glu Met Ile Ser Ser Gly Val Tyr Pro Asp Thr Ile Thr Ile Arg Asn  
 645                      650                      655  
 atg ctg act ggt tta tgg agt aaa gag gaa cta aaa agg gca gtg gca 2016  
 Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Ala Val Ala  
 660                      665                      670  
 atg ctt gag aaa ctg cag atg agt atg gat cta tca ttt ggg gga tga 2064  
 Met Leu Glu Lys Leu Gln Met Ser Met Asp Leu Ser Phe Gly Gly Xaa  
 675                      680                      685  
 <210> 17  
 <211> 688  
 <212> PRT  
 <213> Raphanus sativus  
 <400> 17  
 Met Leu Ala Arg Val Cys Gly Phe Lys Cys Ser Ser Ser Pro Ala Glu  
 1                      5                      10                      15  
 Ser Ala Ala Arg Leu Phe Cys Thr Arg Ser Ile Arg Asp Thr Leu Ala  
 20                      25                      30  
 Lys Ala Ser Gly Glu Ser Cys Glu Ala Gly Phe Gly Gly Glu Ser Leu  
 35                      40                      45

Lys Leu Gln Ser Gly Phe His Glu Ile Lys Gly Leu Glu Asp Ala Ile  
 50 55 60  
 Asp Leu Phe Ser Asp Met Leu Arg Ser Arg Pro Leu Pro Ser Val Val  
 65 70 75 80  
 Asp Phe Cys Lys Leu Met Gly Val Val Val Arg Met Glu Arg Pro Asp  
 85 90 95  
 Leu Val Ile Ser Leu Tyr Gln Lys Met Glu Arg Lys Gln Ile Arg Cys  
 100 105 110  
 Asp Ile Tyr Ser Phe Asn Ile Leu Ile Lys Cys Phe Cys Ser Cys Ser  
 115 120 125  
 Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys Leu Thr Lys Leu Gly  
 130 135 140  
 Leu His Pro Asp Val Val Thr Phe Thr Thr Leu Leu His Gly Leu Cys  
 145 150 155 160  
 Val Glu Asp Arg Val Ser Glu Ala Leu Asn Leu Phe His Gln Met Phe  
 165 170 175  
 Glu Thr Thr Cys Arg Pro Asn Val Val Thr Phe Thr Thr Leu Met Asn  
 180 185 190  
 Gly Leu Cys Arg Glu Gly Arg Ile Val Glu Ala Val Ala Leu Leu Asp  
 195 200 205  
 Arg Met Met Glu Asp Gly Leu Gln Pro Thr Gln Ile Thr Tyr Gly Thr  
 210 215 220  
 Ile Val Asp Gly Met Cys Lys Lys Gly Asp Thr Val Ser Ala Leu Asn  
 225 230 235 240  
 Leu Leu Arg Lys Met Glu Glu Val Ser His Ile Ile Pro Asn Val Val  
 245 250 255  
 Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg His Ser

260	265	270
Asp Ala Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile Phe Pro		
275	280	285
Asp Leu Phe Thr Tyr Asn Ser Met Ile Val Gly Phe Cys Ser Ser Gly		
290	295	300
Arg Trp Ser Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu Arg Lys		
305	310	315
Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala Phe Val		
325	330	335
Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu Met Leu		
340	345	350
Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met Ile Asp		
355	360	365
Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met Phe Tyr		
370	375	380
Leu Met Ala Thr Lys Gly Cys Ser Pro Asn Leu Ile Thr Phe Asn Thr		
385	390	395
Leu Ile Asp Gly Tyr Cys Gly Ala Lys Arg Ile Asp Asp Gly Met Glu		
405	410	415
Leu Leu His Glu Met Thr Glu Thr Gly Leu Val Ala Asp Thr Thr Thr		
420	425	430
Tyr Asn Thr Leu Ile His Gly Phe Tyr Leu Val Gly Asp Leu Asn Ala		
435	440	445
Ala Leu Asp Leu Leu Gln Glu Met Ile Ser Ser Gly Leu Cys Pro Asp		
450	455	460
Ile Val Thr Cys Asp Thr Leu Leu Asp Gly Leu Cys Asp Asn Gly Lys		
465	470	475
		480

Leu Lys Asp Ala Leu Glu Met Phe Lys Val Met Gln Lys Ser Lys Lys  
 485 490 495  
 Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp Val Gln  
 500 505 510  
 Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys Phe Leu  
 515 520 525  
 Glu Ala Glu Glu Leu Tyr Glu Glu Met Pro His Arg Gly Ile Val Pro  
 530 535 540  
 Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys Gln Ser  
 545 550 555 560  
 Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser Lys Ser  
 565 570 575  
 Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Asn Gly Tyr Cys  
 580 585 590  
 Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu Met Gly  
 595 600 605  
 Arg Arg Gly Ile Val Ala Asn Ala Ile Thr Tyr Ile Thr Leu Ile Cys  
 610 615 620  
 Gly Phe Arg Lys Val Gly Asn Ile Asn Gly Ala Leu Asp Ile Phe Gln  
 625 630 635 640  
 Glu Met Ile Ser Ser Gly Val Tyr Pro Asp Thr Ile Thr Ile Arg Asn  
 645 650 655  
 Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Ala Val Ala  
 660 665 670  
 Met Leu Glu Lys Leu Gln Met Ser Met Asp Leu Ser Phe Gly Gly Xaa  
 675 680 685

&lt;210&gt; 18

&lt;211&gt; 2073

&lt;212&gt; DNA

<213> *Raphanus sativus*

&lt;400&gt; 18

atg ttg gct agg gtt tgt gga ttc aag tgt tct tct cct gct gtg 48

Met Leu Ala Arg Val Cys Gly Phe Lys Cys Ser Ser Ser Pro Ala Val

1

5

10

15

tct gcg gct aga ttg ttc tgt acg aga tog att cgt gat act ctg gcc 96

Ser Ala Ala Arg Leu Phe Cys Thr Arg Ser Ile Arg Asp Thr Leu Ala

20

25

30

aag gca agc agg gat gga gag agt tgc gaa gca ggt ttt gga gga gag 144

Lys Ala Ser Arg Asp Gly Glu Ser Cys Glu Ala Gly Phe Gly Gly Glu

35

40

45

agt ttg aag ctg caa agt ggg ttt cat gaa atc aaa ggt tta gag gat 192

Ser Leu Lys Leu Gln Ser Gly Phe His Glu Ile Lys Gly Leu Glu Asp

50

55

60

gcg att gat ttg ttc agt gac atg ctt cga tct cgt cct tta cct tct 240

Ala Ile Asp Leu Phe Ser Asp Met Leu Arg Ser Arg Pro Leu Pro Ser

65

70

75

80

gtg gtt gat ttc tgt aaa ttg atg ggt gtg gtg gtg agg atg aaa cgc 288

Val Val Asp Phe Cys Lys Leu Met Gly Val Val Val Arg Met Lys Arg

85

90

95

ccg gat gtt gtg att tct ctc cat aag aag atg gaa atg cgg cgc att 336

Pro Asp Val Val Ile Ser Leu His Lys Lys Met Glu Met Arg Arg Ile

100

105

110

cca tgt gat gca tac agc ttc aat att ctg ata aag tgt ttc tgc agc 384

Pro Cys Asp Ala Tyr Ser Phe Asn Ile Leu Ile Lys Cys Phe Cys Ser

115	120	125	
tgc tct aag ctg ccc ttt gct ttg tct aca ttt ggt aag ctc acc aag	432		
Cys Ser Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys Leu Thr Lys			
130	135	140	
ctt gga ctc cac cct gat gtt gtt acc ttc acc acc ctt ctc cac gga	480		
Leu Gly Leu His Pro Asp Val Val Thr Phe Thr Thr Leu Leu His Gly			
145	150	155	160
ttg tgt gtg gaa aat agg ggt tct gaa gct ttg aat ttg ttt cat caa	528		
Leu Cys Val Glu Asn Arg Gly Ser Glu Ala Leu Asn Leu Phe His Gln			
165	170	175	
atg ttt gaa acg caa tgt agg ccc aat gtc gta acc ttc acc act ttg	576		
Met Phe Glu Thr Thr Cys Arg Pro Asn Val Val Thr Phe Thr Thr Leu			
180	185	190	
atg aac ggt ctt tgc cgc gag ggt aga att gtc gaa gcc gta gct cta	624		
Met Asn Gly Leu Cys Arg Glu Gly Arg Ile Val Glu Ala Val Ala Leu			
195	200	205	
ctt gat egg atg atg gaa gat ggt ctc cag cct acc cag att act tat	672		
Leu Asp Arg Met Met Glu Asp Gly Leu Gln Pro Thr Gln Ile Thr Tyr			
210	215	220	
gga aca atc gta gat ggg atg tgt aag aag gga gat act gtg tct gca	720		
Gly Thr Ile Val Asp Gly Met Cys Lys Lys Gly Asp Thr Val Ser Ala			
225	230	235	240
ctg aat ctg ctg agg aag atg gag gag gtg agc cac atc ata ccc aat	768		
Leu Asn Leu Leu Arg Lys Met Glu Glu Val Ser His Ile Ile Pro Asn			
245	250	255	
gtt gta atc tat agt gca atc att gat agc ctt tgt aaa gac gga cgt	816		
Val Val Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg			

260	265	270	
cat agc gat tct caa aat ctt ttc act gaa atg caa gag aaa gga atc			864
His Ser Asp Ser Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile			
275	280	285	
ttt cca gat tta ttt acc tac aac tgt atg atc aac ggg ttt tgt agc			912
Phe Pro Asp Leu Phe Thr Tyr Asn Cys Met Ile Asn Gly Phe Cys Ser			
290	295	300	
tct ggt aga tgg atc gac gcg gag cag ttg ttg caa gaa atg tta gaa			960
Ser Gly Arg Trp Ile Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu			
305	310	315	320
agg aag atc agc cct gat gtt gta act tat aat gct ttg atc aat gaa			1008
Arg Lys Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala			
325	330	335	
ttt gtc aag gaa ggc aag ttc ttt gag gct gaa gaa tta tao gat gag			1056
Phe Val Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu			
340	345	350	
atg ctt cct agg ggt ata atc cct aat aca atc aca tat agt tca atg			1104
Met Leu Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met			
355	360	365	
atc gat gga ttt tgc aaa cag aat cgt ctt gat gct gct gag cac atg			1152
Ile Asp Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met			
370	375	380	
ttt tat ttg atg cct acc aag ggc tgc tct cag gac gta ttc act ttc			1200
Phe Tyr Leu Met Pro Thr Lys Gly Cys Ser Pro Asp Val Phe Thr Phe			
385	390	395	400
aat act ctc ata gac gga tat cgt ggg gct aag agg ata gat gat gga			1248
Asn Thr Leu Ile Asp Gly Tyr Arg Gly Ala Lys Arg Ile Asp Asp Gly			

405                      410                      415  
 atg gaa ctt ctc cat gag atg act gaa gca gga tta gtt gct aac aca 1296  
 Met Glu Leu Leu His Glu Met Thr Glu Ala Gly Leu Val Ala Asn Thr  
 420                      425                      430  
 gtt act tac aac act ctt att cac ggg ttt tgt cag gtg ggc gat ctt 1344  
 Val Thr Tyr Asn Thr Leu Ile His Gly Phe Cys Gln Val Gly Asp Leu  
 435                      440                      445  
 act gct gct cta gac ctt cta cat gag atg att tct agt ggt gtg tgc 1392  
 Thr Ala Ala Leu Asp Leu Leu His Glu Met Ile Ser Ser Gly Val Cys  
 450                      455                      460  
 cct aat gtc gtt act tgt agc act ttg ctg gat ggt ctc tgc gat aac 1440  
 Pro Asn Val Val Thr Cys Ser Thr Leu Leu Asp Gly Leu Cys Asp Asn  
 465                      470                      475                      480  
 ggg aaa cta aaa gat gca tgg gua ctg ttt aag gtt atg cag aag agt 1488  
 Gly Lys Leu Lys Asp Ala Trp Glu Leu Phe Lys Val Met Gln Lys Ser  
 485                      490                      495  
 aag atg gat ctt gat gct agt cac ccc ttc aat ggt gtg gaa cct gat 1536  
 Lys Met Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp  
 500                      505                      510  
 gtt caa act tac aat ata ttg atc agc ggc ttg atc aat gaa ggg aag 1584  
 Val Gln Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys  
 515                      520                      525  
 ttt tta gag gct gag gaa tta tac aag gag atg ccc cac agg ggt ata 1632  
 Phe Leu Glu Ala Glu Glu Leu Tyr Lys Glu Met Pro His Arg Gly Ile  
 530                      535                      540  
 gtc cca gat act att acc tat agc tca atg atc gat gga cta tgc aag 1680  
 Val Pro Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys

545	550	555	560
cag agc cgc ctg gat gag gct aca caa atg ttt gat teg atg ggt agc 1728			
Gln Ser Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser			
565	570	575	
aag agc ttc tct cca aac gta gtg acc ttt act aca ctc att gat ggc 1776			
Lys Ser Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Asp Gly			
580	585	590	
tac tgt aaa gca gga agg gtt gat gat ggg ctg gag ctt ttc tgc gag 1824			
Tyr Cys Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu			
595	600	605	
atg ggt aga aga ggg ata gtt gct aat aca att act tac atc act ttg 1872			
Met Gly Arg Arg Gly Ile Val Ala Asn Thr Ile Thr Tyr Ile Thr Leu			
610	615	620	
att cgt ggt ttt cgc aat gtg ggt aat att aat ggg gct cta gac att 1920			
Ile Arg Gly Phe Arg Asn Val Gly Asn Ile Asn Gly Ala Leu Asp Ile			
625	630	635	640
ttc cag gag atg att tca agt ggt gtg tat cct ggt atc att act atc 1968			
Phe Gln Glu Met Ile Ser Ser Gly Val Tyr Pro Gly Ile Ile Thr Ile			
645	650	655	
cgc agt atg ctg act ggt tta tgg agt aaa gag gaa cta aaa agg aca 2016			
Arg Ser Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Thr			
660	665	670	
gtg gca atg ctt gag gaa ctg cag atg agt gtg ggg tat cag ttg gag 2064			
Val Ala Met Leu Glu Glu Leu Gln Met Ser Val Gly Tyr Gln Leu Glu			
675	680	685	
gat gaa tga			2073
Asp Glu Xaa			

545	550	555	560
cag agc cgc ctg gat gag gct aca caa atg ttt gat tcg atg ggt agc 1728			
Gln Ser Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser			
565	570	575	
aag agc ttc tet cca aac gta gtg acc ttt act aca ctc att gat ggc 1776			
Lys Ser Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Asp Gly			
580	585	590	
tac tgt aaa gca gga agg gtt gat gat ggg ctg gag ctt ttc tgc gag 1824			
Tyr Cys Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu			
595	600	605	
atg ggt aga aga ggg ata gtt gct aat aca att act tac atc act ttg 1872			
Met Gly Arg Arg Gly Ile Val Ala Asn Thr Ile Thr Tyr Ile Thr Leu			
610	615	620	
att cgt ggt ttt cgc aat gtg ggt aat att aat ggg gct cta gac att 1920			
Ile Arg Gly Phe Arg Asn Val Gly Asn Ile Asn Gly Ala Leu Asp Ile			
625	630	635	640
ttc cag gag atg att tca agt ggt gtg tat cct ggt atc att act atc 1968			
Phe Gln Glu Met Ile Ser Ser Gly Val Tyr Pro Gly Ile Ile Thr Ile			
645	650	655	
cgc agt atg ctg act ggt tta tgg agt aaa gag gaa cta aaa agg aca 2016			
Arg Ser Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Thr			
660	665	670	
gtg gca atg ctt gag gaa ctg cag atg agt gtg ggg tat cag ttg gag 2064			
Val Ala Met Leu Glu Glu Leu Gln Met Ser Val Gly Tyr Gln Leu Glu			
675	680	685	
gat gaa tga			2073
Asp Glu Xaa			

	165	170	175
Met Phe Glu Thr Thr Cys Arg Pro Asn Val Val Thr Phe Thr Thr Leu			
180	185	190	
Met Asn Gly Leu Cys Arg Glu Gly Arg Ile Val Glu Ala Val Ala Leu			
195	200	205	
Leu Asp Arg Met Met Glu Asp Gly Leu Gln Pro Thr Gln Ile Thr Tyr			
210	215	220	
Gly Thr Ile Val Asp Gly Met Cys Lys Lys Gly Asp Thr Val Ser Ala			
225	230	235	240
Leu Asn Leu Leu Arg Lys Met Glu Glu Val Ser His Ile Ile Pro Asn			
245	250	255	
Val Val Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg			
260	265	270	
His Ser Asp Ser Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile			
275	280	285	
Phe Pro Asp Leu Phe Thr Tyr Asn Cys Met Ile Asn Gly Phe Cys Ser			
290	295	300	
Ser Gly Arg Trp Ile Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu			
305	310	315	320
Arg Lys Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala			
325	330	335	
Phe Val Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu			
340	345	350	
Met Leu Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met			
355	360	365	
Ile Asp Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met			
370	375	380	

Phe Tyr Leu Met Pro Thr Lys Gly Cys Ser Pro Asp Val Phe Thr Phe  
 385 390 395 400  
 Asn Thr Leu Ile Asp Gly Tyr Arg Gly Ala Lys Arg Ile Asp Asp Gly  
 405 410 415  
 Met Glu Leu Leu His Glu Met Thr Glu Ala Gly Leu Val Ala Asn Thr  
 420 425 430  
 Val Thr Tyr Asn Thr Leu Ile His Gly Phe Cys Gln Val Gly Asp Leu  
 435 440 445  
 Thr Ala Ala Leu Asp Leu Leu His Glu Met Ile Ser Ser Gly Val Cys  
 450 455 460  
 Pro Asn Val Val Thr Cys Ser Thr Leu Leu Asp Gly Leu Cys Asp Asn  
 465 470 475 480  
 Gly Lys Leu Lys Asp Ala Trp Glu Leu Phe Lys Val Met Gln Lys Ser  
 485 490 495  
 Lys Met Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp  
 500 505 510  
 Val Gln Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys  
 515 520 525  
 Phe Leu Glu Ala Glu Glu Leu Tyr Lys Glu Met Pro His Arg Gly Ile  
 530 535 540  
 Val Pro Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys  
 545 550 555 560  
 Gln Ser Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser  
 565 570 575  
 Lys Ser Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Asp Gly  
 580 585 590  
 Tyr Cys Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu

595                      600                      605  
 Met Gly Arg Arg Gly Ile Val Ala Asn Thr Ile Thr Tyr Ile Thr Leu  
 610                      615                      620  
 Ile Arg Gly Phe Arg Asn Val Gly Asn Ile Asn Gly Ala Leu Asp Ile  
 625                      630                      635                      640  
 Phe Gln Glu Met Ile Ser Ser Gly Val Tyr Pro Gly Ile Ile Thr Ile  
 645                      650                      655  
 Arg Ser Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Thr  
 660                      665                      670  
 Val Ala Met Leu Glu Glu Leu Gln Met Ser Val Gly Tyr Gln Leu Glu  
 675                      680                      685  
 Asp Glu Xaa  
 690  
 <210> 20  
 <211> 516  
 <212> DNA  
 <213> *Raphanus raphanistrum*  
 <400> 20  
 a atg gaa cgc ccg gat ctt gtg att tet ctc tat caa aag atg gaa agg 49  
 Met Glu Arg Pro Asp Leu Val Ile Ser Leu Tyr Gln Lys Met Glu Arg  
 1                      5                      10                      15  
 aaa cag att cca tgt gat gta tac agc ttt aat att ctg ata aaa tgt 97  
 Lys Gln Ile Pro Cys Asp Val Tyr Ser Phe Asn Ile Leu Ile Lys Cys  
 20                      25                      30  
 ttc tgc agt tgc tct aag ctt ccc ttt gct ttg tct aca ttt ggt aag 145  
 Phe Cys Ser Cys Ser Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys  
 35                      40                      45

```

aie acc aag ctt gga ctc cac cct gat gtt gct acc ttc aac acc ctg 193
Ile Thr Lys Leu Gly Leu His Pro Asp Val Ala Thr Phe Asn Thr Leu

      50              55              60

ctc cac gga tta tgt ctt gat aag agg gtt tct gaa gcc ttg gat ttg 241
Leu His Gly Leu Cys Leu Asp Lys Arg Val Ser Glu Ala Leu Asp Leu

      65              70              75              80

ttt cat caa atg ttt gaa acg aca tgt agg ccg aac atc ata acg ttt 289
Phe His Gln Met Phe Glu Thr Thr Cys Arg Pro Asn Ile Ile Thr Phe

              85              90              95

acc acg ctg atg aac ggt ctt tgc tac gag ggt aga gtt gtc gaa gct 337
Thr Thr Leu Met Asn Gly Leu Cys Tyr Glu Gly Arg Val Val Glu Ala

      100             105             110

gta gct ctg ctt gat cgg atg cta gaa gat ggt ctc cag cct gac cag 385
Val Ala Leu Leu Asp Arg Met Leu Glu Asp Gly Leu Gln Pro Asp Gln

      115             120             125

att act tac gga aca att gta gac ggg atg tgt aag atg gga gac act 433
Ile Thr Tyr Gly Thr Ile Val Asp Gly Met Cys Lys Met Gly Asp Thr

      130             135             140

gtg tct gca ttg aat ctt ctg agg aag atg gag gag tlg agc cac atc 481
Val Ser Ala Leu Asn Leu Leu Arg Lys Met Glu Glu Leu Ser His Ile

      145             150             155             160

aaa ccc aat gtg gta atc tat agt gcc atc att ga 516
Lys Pro Asn Val Val Ile Tyr Ser Ala Ile Ile

      165             170

```

&lt;210&gt; 21

&lt;211&gt; 171

&lt;212&gt; DNA

<213> *Raphanus raphanistrum*

&lt;400&gt; 21

Met Glu Arg Pro Asp Leu Val Ile Ser Leu Tyr Gln Lys Met Glu Arg

1 5 10 15

Lys Gln Ile Pro Cys Asp Val Tyr Ser Phe Asn Ile Leu Ile Lys Cys

20 25 30

Phe Cys Ser Cys Ser Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys

35 40 45

Ile Thr Lys Leu Gly Leu His Pro Asp Val Ala Thr Phe Asn Thr Leu

50 55 60

Leu His Gly Leu Cys Leu Asp Lys Arg Val Ser Glu Ala Leu Asp Leu

65 70 75 80

Phe His Gln Met Phe Glu Thr Thr Cys Arg Pro Asn Ile Ile Thr Phe

85 90 95

Thr Thr Leu Met Asn Gly Leu Cys Tyr Glu Gly Arg Val Val Glu Ala

100 105 110

Val Ala Leu Leu Asp Arg Met Leu Glu Asp Gly Leu Gln Pro Asp Gln

115 120 125

Ile Thr Tyr Gly Thr Ile Val Asp Gly Met Cys Lys Met Gly Asp Thr

130 135 140

Val Ser Ala Leu Asn Leu Leu Arg Lys Met Glu Glu Leu Ser His Ile

145 150 155 160

Lys Pro Asn Val Val Ile Tyr Ser Ala Ile Ile

165 170

&lt;210&gt; 22

&lt;211&gt; 2073

&lt;212&gt; DNA

&lt;213&gt; Raphanus

&lt;400&gt; 22

atgttgctta gsgtttgtag attcaagtgt tcttctcttc ctgcctgwtc tggcgctaga 60  
 ttgttctgta cgagatcgat tctgtatact ctggccaagg caagcgrga krmnnnnngt 120  
 tgcgaagcag gttttgag agagagtttg aagctgraaa gtgggtttca tgaatcaaa 180  
 ggttttagag atcgatgta tttgttcagt gacatgcttc gatctgctcc tttaacctct 240  
 gtggttgatt tctgtaatt gatgggtgtg gtggtagra tgraacgcc gatatgtgtg 300  
 atttctctcy atmaragat ggaakgmrr crsatccsat gtgatyata cagcttyaat 360  
 attctgataa artgtttctg cagytgctct aagcttccct ttgctttgtc tacatttggt 420  
 aagmtacca agcttgagat ccacctgat gttgtyacct tcamaccct kotccayga 480  
 ttrtgystkg awrakaggak ttctgaagcy ttgrattikt tctatcaatt gtttgaacg 540  
 roatgtaggc csaayrtct aacsttyacc ackytgatga acggtctttg cyrcagaggt 600  
 agarttgctg aagcytagc tctcttgat cggatgmtrg aagatggctt ccagcctrmc 660  
 cagatttact ayggaacaa ygtagaygg atgtgtaaga wgggagayac tgtgtctgca 720  
 ytgaatctkc tgaggaaagt ggaggagktg agccacatca waccaatgt kgtaatctat 780  
 agtgcmatca ttgtagcct ttgtaagac ggacgtcata gcatkwcra aaatcttttc 840  
 actgaaatgc aagagaaagg aatctttcom gatttattta cctacaacwg tatgatmwy 900  
 gktttttgta gctctgtag atggakcgac gcggagcagt tgttgcaaga aatgttagaa 960  
 aggaagatca gccctgatgt tgaacttat aatgctttga tcaatgcatt tgtcaaggaa 1020  
 gacaagtctt ttgagctga agaattatac gatgagatgc ttocwagggg tataatccct 1080  
 aatacaatca catatagtic aatgatcgat ggattttgca aacagaaatg tcttgatgct 1140  
 gctgagcaca tgtttatit gatgctacc aagggtgct ctcscract awtactttc 1200  
 aatactctca tagacggata tytgggggct aagaggatag atgatggaat ggaactctc 1260  
 catgagatga ctgaarcagg attagtgtgt racacarya cttaacacac tcttattcac 1320  
 gggtytrtc wgttgggga tcttamtgct gctctagacc tttacawga gatgatyctt 1380  
 agtggktgt gccctratrt cgttacttgt rraactttgc tggatggtct ctgcgataay 1440  
 gggaactcaa aagatgcak ggaamtgttt aaggttatgc agaagagtaa gawgatctt 1500

gatgctagtc acccttcaa tgggtggaa cctgatgttc aaacttacaa tatattgac 1560  
 agcggttga tcaatgaagg gaagttttta gaggcygagg aattatacra ggagatgcc 1620  
 cacagggta tagtccaga tactatyacc tatagctcaa tgatgatgg aytatgcaag 1680  
 cagagccgcc tggatgagc tacacaaatg ttgtattcga tgggtagcaa gagcttctct 1740  
 ccaaaagtag tgacatttac tacacicatt ratggctact gtaargcagg aagggttgat 1800  
 gatgggttg agctttttctg cgagatgggt maaagaggga tagtigctaa yrcaattact 1860  
 tacatcatt tgatttgtag ttttcgyaaw ggggtaata ttaatgggc tctagacatt 1920  
 ttccaggaga tgatttcaag tgggtgtat cctgrtayca ttacyatccg cartatgtg 1980  
 actggtttat ggagtaaaga ggaactaana aggrecagtg caatgcttga graactgcag 2040  
 atgagtrtg rkywwyartt kgrgrwkra tga 2073

&lt;210&gt; 23

&lt;211&gt; 2073

&lt;212&gt; DNA

&lt;213&gt; Raphanus

&lt;400&gt; 23

atgttggcta gggtttgagg attcaagtg tcttcttctc ctgctgwtc tgcggtaga 60  
 ttgtttctga cgagatgat tegtatact ctggccaagg caagcgrga kmnnnnngt 120  
 tgcgaagcag gttttggagg agagagttg aagctgcana gtgggttca tgaatcaaa 180  
 ggtttagagg atgcgaltga tttgttcagt gacatgctc gatctgtcc ttacortct 240  
 gtggttgatt totgtaaat gatgggtgig gtggtgagra tgraacgcc gzatsttg 300  
 atttctctcy atmagaagt ggaakgmrr crsatccsat gtatryata cagcttcaat 360  
 attctgataa artgtttctg cagctgtctt aagctaccct ttgctttgtc tacatttggt 420  
 aagmtacca agcttggaat ccacctgat gttgttacct tccaccacct kotccaygga 480  
 ttrtygttg aarataaggk tctgaagcy ttgrattkt ttcataaat gtttgaacg 540  
 rcattgtagc ccaatgtcgt aaccttcaac accttgnatg acggtctttg ccgcagagg 600  
 agaattgtcg aagccgtagc tctrettgat cggatgatgg aagatggtt ccagcttacc 660  
 cagattactt atggaacaat cgtagaiggg atgigtgaga agggagatac tgtgtctgca 720

ctgaatctgc tgggaagat ggaggaggig agccacatca taccgaatgt tgaatctat 780  
 agtgcacatc ttgatagcct ttgtaagac ggaagtcata gcatkewca aatctcttct 840  
 actgaatgc aagagaagg aatcttccm gatttatita cctacaacwg tatgatwry 900  
 ggkttttgia gctctggtag atggakcgac gggagcagt tgttgcaaga aatgttagaa 960  
 aggaagatca gccctgatgt tgaacttat aatgcitiga tcaatgcatt tgcgaaggaa 1020  
 ggcaagtctt ttgaggctga agaattatac gatgagatgc ttccwagggg tataatccct 1080  
 aatacacatc catatagtic aatgatcgat ggatttttga aacagaaatg tcttgatgt 1140  
 gctgagocaa tgttttatt gatgctacc aagggtctgt ctccraact awtcaacttc 1200  
 aatactctca tagacggata tytgaggct aagaggatag atgatggast ggaactcttc 1260  
 catgagatga ctgaarcagg attagtgtct racacarya cttacacac tottattcac 1320  
 gggttytrtc wggtagggca tctamtgtct gctctagacc tttacawga gatgatytct 1380  
 agtggtktgt gccctratrt cgttacttgt rrcactllgc tggatggtct ctgcgataay 1440  
 gggaactaa aagatgcath ggaamtgttt aaggttatgc agaagagtaa gawgatctt 1500  
 gatgctatgc acccctcaa tgggttgga cctgatgttc aaacttacaa tatattgac 1560  
 agcggcttga tcaatgaagg gaagttttta gaggcygagg aattatacra ggagatgccc 1620  
 cacagggga tagtccaga tactatyacc tatagctcaa tgatcgatgg aytatgcaag 1680  
 cagagccgcc tggatgagcg tacacaaatg tttgattcga tgggttagcaa gagcttctct 1740  
 ccanaagtag tgacctttac tacaciatc ratggctact gtaargcagg aagggttgat 1800  
 gatggctgag agctttttct cgagatgggt msaagaggga tagttgctaa yrcaattact 1860  
 tacaacactt tgatttygtg ttttcgaaaw gtgggtaata ttaatggggc tctagacatt 1920  
 ttccaggaga tgatttcaag tgggtgigtat cctgriayca ttacyatccg cartatgctg 1980  
 acttggttat ggagttaaga ggaactaaa aggrcagtg caatgcttga gnaactgcag 2040  
 atgagtrigg rkywwymrtt kgrggrwkra tga 2073

&lt;210&gt; 24

&lt;211&gt; 2064

&lt;212&gt; DNA

&lt;213&gt; Raphanus

&lt;400&gt; 24

atgttggota gggtttgigs attcaagigt tcttcttctc ctgctgngtc tggggetaga 60  
 ttgtttctgta cagagatcgat tctgtatact ctggccaagg caagcggaga gattttgcgaa 120  
 gcaggttttg gaggagagag tttnaagctg caaagtgggt ttcattgaaat caaagggtta 180  
 gaggatcgga ttgatttggt cagtacatg ctctgatctc gtcctttacc ttctgttggt 240  
 gattttctgta aattgatggg tgtgttggtg agaattggaac gcccggtatct tctgatttct 300  
 ctctatcara agatggaaaag gaaacagatt csatgtgatr tatacagctt yaattattctg 360  
 ataaaatggt totgcagytg ctctaagcty ccctttgctt tgtctacatt tggtaagmtc 420  
 accaagcttg gactccacco tcatgttgyt acccttcamca cctgtctoca yggatttrgy 480  
 stkgawraka gggtttctga agcyttgrat tkttttcttc aaagtttga aacgacatgt 540  
 aggccsaayr tertaacstt yaocackytg atgaacggte ttgcyroga gggtagartt 600  
 gtctgaagcyg tagctctgct tgatoggatg mtrgaagatg gtctccagcc trmcagatt 660  
 acttaygaa caatygtaga ygggatgtgt aagawggag ayactgtctc tgcaytgaat 720  
 ctktctgaga agatggagga gktgagccac atcawacca atgktgtaat ctatagtcm 780  
 atcattgata gcttttgtaa agacggacgt catagcgtg cccaaaatct tttaactgaa 840  
 atgcaagaga aaggaatctt tccgattha ttacctaca acagtatgat agttggtttt 900  
 tctagctctg gtagatggag cgacggggag cagttgttc aggaatggt agaaaggaaag 960  
 atcagccctg atgttgttaa ttataatgot ttgatcaatg catttgctaa ggaaggcaag 1020  
 ttctttgagg ctgaagaatt atacgatgag atgcttccaa ggggtataat cctaataca 1080  
 atcacatata gtccaatgat cgatggattt tgcaaacaga atcgtcttga tctgtctgag 1140  
 cacatgtttt atttgaigc taccnagggc tgcctfcca acctastcac ttcaatact 1200  
 ctcatagacg gatattgttg ggcctnaggg atagatgatg gaatggaact tctcatgag 1260  
 atgaactgaa caggattagt tctgacaca actacttaca acactcttat tcacgggttc 1320  
 tctctgttg gogactctaa tctgtctcta gaacttttac aagagatgat ctctagtgtt 1380  
 ttgtgccctg atatcgttac ttgtgacact ttgtgtggat gtcctctcga taattggaaa 1440  
 ctanaagatg cattggaat gttaaagtt atgcagaaga gtaagaagga tcttgatgct 1500  
 agtcacectt tcaatgggtt ggaacctgat gtccaactt acaatatatt gatcagcgc 1560

ttgatcaatg aagggaagtt ttagaggcc gaggaattat acgaggagat gccccacagg 1620  
 ggtatagtcc cagaaactat caactatagc tcaatgatcg atggattatg caagcagagc 1680  
 cgcctagatg aggtacaca aatgtttgat tcgatgggta gcaagagctt ctctccaaac 1740  
 gtagtgacct ttactacact cattaatggc tactgttaagg caggaaaggt tgatgatggg 1800  
 ctggagcttt tctgcgagat gggtogaaga gggatagtfg ctaacgcaat tacttacatc 1860  
 accttgattt gtgggtttcg taaagtgggt aatattaatg gggctctaga catittccag 1920  
 gagatgattt caagtggtgt gtatcctgat accattacca tccgcaatat gctgactggt 1980  
 ttatggagta aagaggaact aaaaagggca gtggcaatgc ttgagaaact gcagatgagt 2040  
 atggatctat catttggggg atga 2064

<210> 25

<211> 2064

<212> DNA

<213> Raphanus

<400> 25

atgttggcta gggtttggg attcaagtgt tcttctctc ctgctgagtc tgcggctaga 60  
 ttgttctgta cagagatcat tctgatact ctggccaagg caagcggaga gagttcgaa 120  
 gcaggttttg gaggcgagag ttgaagctg caaagtgggt tcatgaaat caaaggttta 180  
 gaggatcgca ttgatitgt cagtgcacatg ctctgatctc gtcccttacc ttctgtgggt 240  
 gatttctgta aattgatggg tgtgggtggg agaattggaac gcccggaatc tgtgatttct 300  
 ctctatcaga agatggaag gaacagatt cgaatgata tatacagctt caatattctg 360  
 ataaaatgtt tctgcagctg ctctaaagtc ccttttgett tgtctacatt tgtaagmtc 420  
 accaagottg gactccaacc tgatgttgtt accttcacca cctgtctcca ygatrttgy 480  
 gtggagata gggtttctga agcyttgrat tiktttctc aantgtttga aacgacatgt 540  
 agggccaatg tcttaacctt caacactttg atgaacgctc ttgcccga gggtagaatt 600  
 gtogaagcgc tagctctgct tgatcgagtg atggaagatg gtctccagcc taccagatt 660  
 acctatggaa caatctaga igggatgtgt aagaaggag atactgtgc tgcactgaat 720  
 ctgctgagga agatggagga ggtgagccac atcataccca atgttghtat ctatagtga 780

atcattgata gcctttgtaa agaaggacgt catagcgatg cacanaatct tttactgaa 840  
 atgcaagaga aaggaaatct tcccgattta ttacattaca acagiatgat agttgggtti 900  
 tgtagctctg gtagatggag cgaaggaggag cagtgtgtgc aagaaatgtt agnaagggaag 960  
 atcagccctg aigtgtaac ttataatgct ttgatcaatg catttgtcaa ggaaggcnaag 1020  
 ttctttgggg ctgaagaatt atacgatgag atgettccaa ggggtataat ccctaataca 1080  
 atcacatata gticaatgat cgatggattt tgaanaacaga atcgtcttga tgcgtgtgag 1140  
 cacatgtttt atttgatggc taccnaaggc tgcctctcca acctaatcac tttcaatact 1200  
 ctcatagacg gataattgtg ggctaagagg atagatgatg gaatggaact tctccatgag 1260  
 atgactgnaa caggatttagt tgcgtacaca actacttaca acactottat tcacgggttc 1320  
 tatctgggtg ccatctttaa tgcgtctcta gaacttttac aagagatgat ctctagtgtt 1380  
 ttgtgcccgt atalegttac ttgtgacact ttgtgggatg gtctctgca taatgggaaa 1440  
 ctlaaagatg cattggaat gttaaagggt atgcagaaga gtaagaagga tottgatgct 1500  
 agtcacccct tcaatgggtt ggaacctgat gtccaactt acaatatatt gatcagcggc 1560  
 ttgalcactg aagggaaggt tttagaggcc gaggaattat acgaggagat gcccncagg 1620  
 ggtatagtc cagatactat cacttatagc tcaatgatgc atggatfatg caagcagagc 1680  
 cgcctagatg aggcacaca aatgtttgat tcgatgggta gcaagagctt ctctccaaa 1740  
 gtagtgacct ttactacact cattaatggc tactgtacgg caggnaagggt tgatgatggg 1800  
 ctggagcttt totgogagat gggtcgaaga gggatagtg ctacgcnaat tacttacatc 1860  
 actttgattt gtggttttcg taaagtgggt aatattaatg gggctctaga cattttccag 1920  
 gagatgattt caagtgtgt gtatcctgat accattacca tcgcnaatat gctgactggt 1980  
 ttatggagta aagaggaact aaaaagggca gtggcaatgc ttganaact gcagatgagt 2040  
 atggatctat catttggggg atga 2064

<210> 26

<211> 690

<212> PRT

<213> Raphanus

<220>

<221> Xaa  
<222> 18  
<223> Glu or Val  
<221> Xaa  
<222> 36  
<223> Arg or none  
<221> Xaa  
<223> 37, Asp or no  
<221> Xaa  
<222> 95  
<223> Glu or Lys  
<221> Xaa  
<222> 99  
<223> Leu or Val  
<221> Xaa  
<222> 104  
<223> Tyr or His  
<221> Xaa  
<222> 105  
<223> Gln or Lys  
<221> Xaa  
<222> 109  
<223> Arg or Met  
<221> Xaa  
<222> 110  
<223> Lys or Arg  
<221> Xaa

<222> 111  
<223> Gln or Arg  
<221> Xaa  
<222> 113  
<223> Arg or Pro  
<221> Xaa  
<222> 116  
<223> Ile, Ala or Val  
<221> Xaa  
<222> 142  
<223> Leu or Ile  
<221> Xaa  
<222> 152  
<223> Val or Ala  
<221> Xaa  
<222> 155  
<223> Thr or Asn  
<221> Xaa  
<222> 163  
<223> Val, or Leu  
<221> Xaa  
<222> 164  
<223> Glu or Asp  
<221> Xaa  
<222> 165  
<223> Asp, Asn or Lys  
<221> Xaa

<222> 167  
<223> Val or Gly  
<221> Xaa  
<222> 172  
<223> Asn or Asp  
<221> Xaa  
<222> 173  
<223> Leu or Phe  
<221> Xaa  
<222> 186  
<223> Val or Ile  
<221> Xaa  
<222> 187  
<223> Val or Ile  
<221> Xaa  
<222> 198  
<223> Arg or Tyr  
<221> Xaa  
<222> 202  
<223> Ile or Val  
<221> Xaa  
<222> 213  
<223> Met or Leu  
<221> Xaa  
<222> 220  
<223> Thr or Asp  
<221> Xaa

<222> 234  
<223> Lys or Met  
<221> Xaa  
<222> 250  
<223> Val or Leu  
<221> Xaa  
<222> 254  
<223> Ile or Lys  
<221> Xaa  
<222> 276  
<223> Ala or Ser  
<221> Xaa  
<222> 297  
<223> Ser or Cys  
<221> Xaa  
<222> 300  
<223> Val or Asn  
<221> Xaa  
<222> 309  
<223> Ser or Ile  
<221> Xaa  
<222> 389  
<223> Ala or Pro  
<221> Xaa  
<222> 396  
<223> Asn or Asp  
<221> Xaa

<222> 397  
<223> Leu or Val  
<221> Xaa  
<222> 398  
<223> Ile or Phe  
<221> Xaa  
<222> 408  
<223> Cys or Arg  
<221> Xaa  
<222> 426  
<223> Thr or Ala  
<221> Xaa  
<222> 431  
<223> Asp or Asn  
<221> Xaa  
<222> 433  
<223> Thr or Val  
<221> Xaa  
<222> 443  
<223> Tyr or Cys  
<221> Xaa  
<222> 444  
<223> Leu or Gln  
<221> Xaa  
<222> 449  
<223> Asn or Thr  
<221> Xaa

<222> 456  
<223> Gln or His  
<221> Xaa  
<222> 463  
<223> Leu or Val  
<221> Xaa  
<222> 466  
<223> Asp or Asn  
<221> Xaa  
<222> 467  
<223> Ile or Val  
<221> Xaa  
<222> 471  
<223> Asp or Ser  
<221> Xaa  
<222> 487  
<223> Leu or Trp  
<221> Xaa  
<222> 489  
<223> Met or Leu  
<221> Xaa  
<222> 498  
<223> Lys or Met  
<221> Xaa  
<222> 537  
<223> Glu or Lys  
<221> Xaa

<222> 591  
<223> Asn or Asp  
<221> Xaa  
<222> 618  
<223> Ala or Thr  
<221> Xaa  
<222> 626  
<223> Cys or Arg  
<221> Xaa  
<222> 630  
<223> Lys or Asn  
<221> Xaa  
<222> 652  
<223> Asp or Gly  
<221> Xaa  
<222> 653  
<223> Thr or Ile  
<221> Xaa  
<222> 658  
<223> Asn or Ser  
<221> Xaa  
<222> 672  
<223> Ala or Thr  
<221> Xaa  
<222> 678  
<223> Lys or Glu  
<221> Xaa

<222> 683  
 <223> Met or Val  
 <221> Xaa  
 <222> 684  
 <223> Asp or Gly  
 <221> Xaa  
 <222> 685  
 <223> Leu or Tyr  
 <221> Xaa  
 <222> 686  
 <223> Ser or Gln  
 <221> Xaa  
 <222> 687  
 <223> Phe or Leu  
 <221> Xaa  
 <222> 688  
 <223> Gly or Glu  
 <221> Xaa  
 <222> 689  
 <223> Gly or Asp  
 <221> Xaa  
 <222> 690  
 <223> Glu or none  
 <400> 26

Met Leu Ala Arg Val Cys Gly Phe Lys Cys Ser Ser Ser Pro Ala Xaa

1 5 10 15

Ser Ala Ala Arg Leu Phe Cys Thr Arg Ser Ile Arg Asp Thr Leu Ala

	20	25	30
Lys	Ala	Ser	Xaa Xaa Gly Glu Ser Cys Glu Ala Gly Phe Gly Gly Glu
35		40	45
Ser	Leu Lys Leu Glu Ser Gly Phe His Glu Ile Lys Gly Leu Glu Asp		
50		55	60
Ala	Ile Asp Leu Phe Ser Asp Met Leu Arg Ser Arg Pro Leu Pro Ser		
65		70	75
Val	Val Asp Phe Cys Lys Leu Met Gly Val Val Val Arg Met Xaa Arg		
	85	90	95
Pro	Asp Xaa Val Ile Ser Leu Xaa Xaa Lys Met Glu Xaa Xaa Xaa Ile		
100		105	110
Xaa	Cys Asp Xaa Tyr Ser Phe Asn Ile Leu Ile Lys Cys Phe Cys Ser		
115		120	125
Cys	Ser Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys Xaa Thr Lys		
130		135	140
Leu	Gly Leu His Pro Asp Val Xaa Thr Phe Xaa Thr Leu Leu His Gly		
145		150	155
Leu	Cys Xaa Xaa Xaa Arg Xaa Ser Glu Ala Leu Xaa Xaa Phe His Glu		
	165	170	175
Met	Phe Glu Thr Thr Cys Arg Pro Asn Xaa Xaa Thr Phe Thr Thr Leu		
180		185	190
Met	Asn Gly Leu Cys Xaa Glu Gly Arg Xaa Val Glu Ala Val Ala Leu		
195		200	205
Leu	Asp Arg Met Xaa Glu Asp Gly Leu Glu Pro Xaa Gln Ile Thr Tyr		
210		215	220
Gly	Thr Ile Val Asp Gly Met Cys Lys Xaa Gly Asp Thr Val Ser Ala		
225		230	235
			240

Leu Asn Leu Leu Arg Lys Met Glu Glu Xaa Ser His Ile Xaa Pro Asn			
	245	250	255
Val Val Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg			
	260	265	270
His Ser Asp Xaa Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile			
	275	280	285
Phe Pro Asp Leu Phe Thr Tyr Asn Xaa Met Ile Xaa Gly Phe Cys Ser			
	290	295	300
Ser Gly Arg Trp Xaa Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu			
	305	310	315
Arg Lys Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala			
	325	330	335
Phe Val Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu			
	340	345	350
Met Leu Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met			
	355	360	365
Ile Asp Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met			
	370	375	380
Phe Tyr Leu Met Xaa Thr Lys Gly Cys Ser Pro Xaa Xaa Xaa Thr Phe			
	385	390	395
Asn Thr Leu Ile Asp Gly Tyr Xaa Gly Ala Lys Arg Ile Asp Asp Gly			
	405	410	415
Met Glu Leu Leu His Glu Met Thr Glu Xaa Gly Leu Val Ala Xaa Thr			
	420	425	430
Xaa Thr Tyr Asn Thr Leu Ile His Gly Phe Xaa Xaa Val Gly Asp Leu			
	435	440	445
Xaa Ala Ala Leu Asp Leu Leu Xaa Glu Met Ile Ser Ser Gly Xaa Cys			

450	455	460
Pro Xaa Xaa Val Thr Cys Xaa Thr Leu Leu Asp Gly Leu Cys Asp Asn		
465	470	475
Gly Lys Leu Lys Asp Ala Xaa Glu Xaa Phe Lys Val Met Gln Lys Ser		
485	490	495
Lys Xaa Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp		
500	505	510
Val Gln Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys		
515	520	525
Phe Leu Glu Ala Glu Glu Leu Tyr Xaa Glu Met Pro His Arg Gly Ile		
530	535	540
Val Pro Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys		
545	550	555
Gln Ser Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser		
565	570	575
Lys Ser Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Xaa Gly		
580	585	590
Tyr Cys Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu		
595	600	605
Met Gly Arg Arg Gly Ile Val Ala Asn Xaa Ile Thr Tyr Ile Thr Leu		
610	615	620
Ile Xaa Gly Phe Arg Xaa Val Gly Asn Ile Asn Gly Ala Leu Asp Ile		
625	630	635
Phe Gln Glu Met Ile Ser Ser Gly Val Tyr Pro Xaa Xaa Ile Thr Ile		
645	650	655
Arg Xaa Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Xaa		
660	665	670

Val Ala Met Leu Glu Xaa Leu Gln Met Ser Xaa Xaa Xaa Xaa Xaa Xaa

675

680

685

Xaa Xaa

690

<210> 27

<211> 690

<212> PRT

<213> Raphanus

<220>

<221> Xaa

<222> 16

<223> Glu or Val

<221> Xaa

<222> 36

<223> Arg or none

<221> Xaa

<223> 37, Asp or none

<221> Xaa

<222> 95

<223> Glu or Lys

<221> Xaa

<222> 99

<223> leu or Val

<221> Xaa

<222> 104

<223> Tyr or His

<221> Xaa

<222> 105  
<223> Gln or Lys  
<221> Xaa  
<222> 109  
<223> Arg or Met  
<221> Xaa  
<222> 110  
<223> Lys or Arg  
<221> Xaa  
<222> 111  
<223> Gln or Arg  
<221> Xaa  
<222> 113  
<223> Arg or Pro  
<221> Xaa  
<222> 116  
<223> Ile or Ala  
<221> Xaa  
<222> 142  
<223> Leu or Ile  
<221> Xaa  
<222> 165  
<223> Asp or Asn  
<221> Xaa  
<222> 167  
<223> Val or Gly  
<221> Xaa



<222> 398  
<223> Ile or Phe  
<221> Xaa  
<222> 408  
<223> Cys or Arg  
<221> Xaa  
<222> 426  
<223> Thr or Ala  
<221> Xaa  
<222> 431  
<223> Asp or Asn  
<221> Xaa  
<222> 433  
<223> Thr or Val  
<221> Xaa  
<222> 443  
<223> Tyr or Cys  
<221> Xaa  
<222> 444  
<223> Leu or Gln  
<221> Xaa  
<222> 449  
<223> Asn or Thr  
<221> Xaa  
<222> 456  
<223> Gln or His  
<221> Xaa

<222> 463  
<223> Leu or Val  
<221> Xaa  
<222> 466  
<223> Asp or Asn  
<221> Xaa  
<222> 467  
<223> Ile or Val  
<221> Xaa  
<222> 471  
<223> Asp or Ser  
<221> Xaa  
<222> 487  
<223> Leu or Trp  
<221> Xaa  
<222> 489  
<223> Met or Leu  
<221> Xaa  
<222> 498  
<223> Lys or Met  
<221> Xaa  
<222> 537  
<223> Glu or Lys  
<221> Xaa  
<222> 591  
<223> Asn or Asp  
<221> Xaa

<222> 618  
<223> Ala or Thr  
<221> Xaa  
<222> 626  
<223> Cys or Arg  
<221> Xaa  
<222> 630  
<223> Lys or Asn  
<221> Xaa  
<222> 652  
<223> Asp or Gly  
<221> Xaa  
<222> 653  
<223> Thr or Ile  
<221> Xaa  
<222> 658  
<223> Asn or Ser  
<221> Xaa  
<222> 672  
<223> Ala or Thr  
<221> Xaa  
<222> 678  
<223> Lys or Glu  
<221> Xaa  
<222> 683  
<223> Met or Val  
<221> Xaa

<222> 684  
 <223> Asp or Gly  
 <221> Xaa  
 <222> 685  
 <223> Leu or Tyr  
 <221> Xaa  
 <222> 686  
 <223> Ser or Gln  
 <221> Xaa  
 <222> 687  
 <223> Phe or Leu  
 <221> Xaa  
 <222> 688  
 <223> Gly or Glu  
 <221> Xaa  
 <222> 689  
 <223> Gly or Asp  
 <221> Xaa  
 <222> 690  
 <223> Glu or none  
 <400> 27

Met Leu Ala Arg Val Cys Gly Phe Lys Cys Ser Ser Ser Pro Ala Xaa

1 5 10 15

Ser Ala Ala Arg Leu Phe Cys Thr Arg Ser Ile Arg Asp Thr Leu Ala

20 25 30

Lys Ala Ser Xaa Xaa Gly Glu Ser Cys Glu Ala Gly Phe Gly Gly Glu

35 40 45

Ser Leu Lys Leu Gln Ser Gly Phe His Glu Ile Lys Gly Leu Glu Asp  
 50 55 60  
 Ala Ile Asp Leu Phe Ser Asp Met Leu Arg Ser Arg Pro Leu Pro Ser  
 65 70 75 80  
 Val Val Asp Phe Cys Lys Leu Met Gly Val Val Val Arg Met Xaa Arg  
 85 90 95  
 Pro Asp Xaa Val Ile Ser Leu Xaa Xaa Lys Met Glu Xaa Xaa Xaa Ile  
 100 105 110  
 Xaa Cys Asp Xaa Tyr Ser Phe Asn Ile Leu Ile Lys Cys Phe Cys Ser  
 115 120 125  
 Cys Ser Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys Xaa Thr Lys  
 130 135 140  
 Leu Gly Leu His Pro Asp Val Val Thr Phe Thr Thr Leu Leu His Gly  
 145 150 155 160  
 Leu Cys Val Glu Xaa Arg Xaa Ser Glu Ala Leu Xaa Xaa Phe His Gln  
 165 170 175  
 Met Phe Glu Thr Thr Cys Arg Pro Asn Val Val Thr Phe Thr Thr Leu  
 180 185 190  
 Met Asn Gly Leu Cys Arg Glu Gly Arg Ile Val Glu Ala Val Ala Leu  
 195 200 205  
 Leu Asp Arg Met Met Glu Asp Gly Leu Gln Pro Thr Gln Ile Thr Tyr  
 210 215 220  
 Gly Thr Ile Val Asp Gly Met Cys Lys Lys Gly Asp Thr Val Ser Ala  
 225 230 235 240  
 Leu Asn Leu Leu Arg Lys Met Glu Glu Val Ser His Ile Ile Pro Asn  
 245 250 255  
 Val Val Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg

260	265	270	
His Ser Asp Xaa Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile			
275	280	285	
Phe Pro Asp Leu Phe Thr Tyr Asn Xaa Met Ile Xaa Gly Phe Cys Ser			
290	295	300	
Ser Gly Arg Trp Xaa Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu			
305	310	315	320
Arg Lys Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala			
325	330	335	
Phe Val Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu			
340	345	350	
Met Leu Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met			
355	360	365	
Ile Asp Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met			
370	375	380	
Phe Tyr Leu Met Xaa Thr Lys Gly Cys Ser Pro Xaa Xaa Xaa Thr Phe			
385	390	395	400
Asn Thr Leu Ile Asp Gly Tyr Xaa Gly Ala Lys Arg Ile Asp Asp Gly			
405	410	415	
Met Glu Leu Leu His Glu Met Thr Glu Xaa Gly Leu Val Ala Xaa Thr			
420	425	430	
Xaa Thr Tyr Asn Thr Leu Ile His Gly Phe Xaa Xaa Val Gly Asp Leu			
435	440	445	
Xaa Ala Ala Leu Asp Leu Leu Xaa Glu Met Ile Ser Ser Gly Xaa Cys			
450	455	460	
Pro Xaa Xaa Val Thr Cys Xaa Thr Leu Leu Asp Gly Leu Cys Asp Asn			
465	470	475	480

Gly Lys Leu Lys Asp Ala Xaa Glu Xaa Phe Lys Val Met Gln Lys Ser			
485	490	495	
Lys Xaa Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp			
500	505	510	
Val Gln Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys			
515	520	525	
Phe Leu Glu Ala Glu Glu Leu Tyr Xaa Glu Met Pro His Arg Gly Ile			
530	535	540	
Val Pro Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys			
545	550	555	560
Gln Ser Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser			
565	570	575	
Lys Ser Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Xaa Gly			
580	585	590	
Tyr Cys Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu			
595	600	605	
Met Gly Arg Arg Gly Ile Val Ala Asn Xaa Ile Thr Tyr Ile Thr Leu			
610	615	620	
Ile Xaa Gly Phe Arg Xaa Val Gly Asn Ile Asn Gly Ala Leu Asp Ile			
625	630	635	640
Phe Gln Glu Met Ile Ser Ser Gly Val Tyr Pro Xaa Xaa Ile Thr Ile			
645	650	655	
Arg Xaa Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Xaa			
660	665	670	
Val Ala Met Leu Glu Xaa Leu Gln Met Ser Xaa Xaa Xaa Xaa Xaa			
675	680	685	
Xaa Xaa			

690  
<210> 28  
<211> 687  
<212> PRT  
<213> Raphanus  
<220>  
<221> Xaa  
<222> 111  
<223> Arg or Pro  
<221> Xaa  
<222> 114  
<222> Ile or Val  
<221> Xaa  
<222> 140  
<223> Leu or Ile  
<221> Xaa  
<222> 150  
<223> Val or Ala  
<221> Xaa  
<222> 153  
<223> Thr or Asn  
<221> Xaa  
<222> 161  
<223> Val or Leu  
<221> Xaa  
<222> 162  
<223> Glu or Asp

<221> Xaa  
<222> 163  
<223> Asp or Lys  
<221> Xaa  
<222> 170  
<223> Asn or Asp  
<221> Xaa  
<222> 171  
<223> Leu or Phe  
<221> Xaa  
<222> 184  
<223> Val or Ile  
<221> Xaa  
<222> 185  
<223> Val or Ile  
<221> Xaa  
<222> 196  
<223> Arg or Tyr  
<221> Xaa  
<222> 200  
<223> Ile or Val  
<221> Xaa  
<222> 211  
<223> Met or Leu  
<221> Xaa  
<222> 218  
<223> Thr or Asp

<221> Xaa  
 <222> 232  
 <223> Lys or Met  
 <221> Xaa  
 <222> 248  
 <223> Val or Leu  
 <221> Xaa  
 <222> 252  
 <223> Ile or Lys  
 <400> 28

Met	Leu	Ala	Arg	Val	Cys	Gly	Phe	Lys	Cys	Ser	Ser	Ser	Pro	Ala	Glu
1				5					10					15	
Ser	Ala	Ala	Arg	Leu	Phe	Cys	Thr	Arg	Ser	Ile	Arg	Asp	Thr	Leu	Ala
			20					25					30		
Lys	Ala	Ser	Gly	Glu	Ser	Cys	Glu	Ala	Gly	Phe	Gly	Gly	Glu	Ser	Leu
			35				40						45		
Lys	Leu	Gln	Ser	Gly	Phe	His	Glu	Ile	Lys	Gly	Leu	Glu	Asp	Ala	Ile
			50				55				60				
Asp	Leu	Phe	Ser	Asp	Met	Leu	Arg	Ser	Arg	Pro	Leu	Pro	Ser	Val	Val
			65			70				75				80	
Asp	Phe	Cys	Lys	Leu	Met	Gly	Val	Val	Val	Arg	Met	Glu	Arg	Pro	Asp
				85						90				95	
Leu	Val	Ile	Ser	Leu	Tyr	Gln	Lys	Met	Glu	Arg	Lys	Gln	Ile	Xaa	Cys
				100					105				110		
Asp	Xaa	Tyr	Ser	Phe	Asn	Ile	Leu	Ile	Lys	Cys	Phe	Cys	Ser	Cys	Ser
				115				120					125		
Lys	Leu	Pro	Phe	Ala	Leu	Ser	Thr	Phe	Gly	Lys	Xaa	Thr	Lys	Leu	Gly

130	135	140	
Leu His Pro Asp Val Xaa Thr Phe Xaa Thr Leu Leu His Gly Leu Cys			
145	150	155	160
Xaa Xaa Xaa Arg Val Ser Glu Ala Leu Xaa Xaa Phe His Gln Met Phe			
	165	170	175
Glu Thr Thr Cys Arg Pro Asn Xaa Xaa Thr Phe Thr Thr Leu Met Asn			
	180	185	190
Gly Leu Cys Xaa Glu Gly Arg Xaa Val Glu Ala Val Ala Leu Leu Asp			
	195	200	205
Arg Met Xaa Glu Asp Gly Leu Gln Pro Xaa Gln Ile Thr Tyr Gly Thr			
	210	215	220
Ile Val Asp Gly Met Cys Lys Xaa Gly Asp Thr Val Ser Ala Leu Asn			
225	230	235	240
Leu Leu Arg Lys Met Glu Glu Xaa Ser His Ile Xaa Pro Asn Val Val			
	245	250	255
Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg His Ser			
	260	265	270
Asp Ala Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile Phe Pro			
	275	280	285
Asp Leu Phe Thr Tyr Asn Ser Met Ile Val Gly Phe Cys Ser Ser Gly			
	290	295	300
Arg Trp Ser Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu Arg Lys			
305	310	315	320
Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala Phe Val			
	325	330	335
Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu Met Leu			
	340	345	350

Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met Ile Asp  
 355 360 365  
 Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met Phe Tyr  
 370 375 380  
 Leu Met Ala Thr Lys Gly Cys Ser Pro Asn Leu Ile Thr Phe Asn Thr  
 385 390 395 400  
 Leu Ile Asp Gly Tyr Cys Gly Ala Lys Arg Ile Asp Asp Gly Met Glu  
 405 410 415  
 Leu Leu His Glu Met Thr Glu Thr Gly Leu Val Ala Asp Thr Thr Thr  
 420 425 430  
 Tyr Asn Thr Leu Ile His Gly Phe Tyr Leu Val Gly Asp Leu Asn Ala  
 435 440 445  
 Ala Leu Asp Leu Leu Gln Glu Met Ile Ser Ser Gly Leu Cys Pro Asp  
 450 455 460  
 Ile Val Thr Cys Asp Thr Leu Leu Asp Gly Leu Cys Asp Asn Gly Lys  
 465 470 475 480  
 Leu Lys Asp Ala Leu Glu Met Phe Lys Val Met Gln Lys Ser Lys Lys  
 485 490 495  
 Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp Val Gln  
 500 505 510  
 Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys Phe Leu  
 515 520 525  
 Glu Ala Glu Glu Leu Tyr Glu Glu Met Pro His Arg Gly Ile Val Pro  
 530 535 540  
 Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys Gln Ser  
 545 550 555 560  
 Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser Lys Ser

	565	570	575
Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Asn Gly Tyr Cys			
	580	585	590
Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu Met Gly			
	595	600	605
Arg Arg Gly Ile Val Ala Asn Ala Ile Thr Tyr Ile Thr Leu Ile Cys			
	610	615	620
Gly Phe Arg Lys Val Gly Asn Ile Asn Gly Ala Leu Asp Ile Phe Gln			
625	630	635	640
Glu Met Ile Ser Ser Gly Val Tyr Pro Asp Thr Ile Thr Ile Arg Asn			
	645	650	655
Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Ala Val Ala			
	660	665	670
Met Leu Glu Lys Leu Gln Met Ser Met Asp Leu Ser Phe Gly Gly			
	675	680	685

<210> 29  
 <211> 687  
 <212> PRT  
 <213> Raphanus  
 <220>  
 <221> Xaa  
 <222> 140  
 <223> Leu or Ile  
 <221> Xaa  
 <222> 170  
 <223> Asn or Asp  
 <221> Xaa

&lt;222&gt; 171

&lt;223&gt; Leu or Phe

&lt;400&gt; 29

```

Met Leu Ala Arg Val Cys Gly Phe Lys Cys Ser Ser Ser Pro Ala Glu
  1             5             10             15
Ser Ala Ala Arg Leu Phe Cys Thr Arg Ser Ile Arg Asp Thr Leu Ala
          20             25             30
Lys Ala Ser Gly Glu Ser Cys Glu Ala Gly Phe Gly Gly Glu Ser Leu
          35             40             45
Lys Leu Gln Ser Gly Phe His Glu Ile Lys Gly Leu Glu Asp Ala Ile
          50             55             60
Asp Leu Phe Ser Asp Met Leu Arg Ser Arg Pro Leu Pro Ser Val Val
          65             70             75             80
Asp Phe Cys Lys Leu Met Gly Val Val Val Arg Met Glu Arg Pro Asp
          85             90             95
Leu Val Ile Ser Leu Tyr Gln Lys Met Glu Arg Lys Gln Ile Arg Cys
          100            105            110
Asp Ile Tyr Ser Phe Asn Ile Leu Ile Lys Cys Phe Cys Ser Cys Ser
          115            120            125
Lys Leu Pro Phe Ala Leu Ser Thr Phe Gly Lys Xaa Thr Lys Leu Gly
          130            135            140
Leu His Pro Asp Val Val Thr Phe Thr Thr Leu Leu His Gly Leu Cys
          145            150            155            160
Val Glu Asp Arg Val Ser Glu Ala Leu Xaa Xaa Phe His Gln Met Phe
          165            170            175
Glu Thr Thr Cys Arg Pro Asn Val Val Thr Phe Thr Thr Leu Met Asn
          180            185            190

```

Gly Leu Cys Arg Glu Gly Arg Ile Val Glu Ala Val Ala Leu Leu Asp  
 195 200 205  
 Arg Met Met Glu Asp Gly Leu Gln Pro Thr Gln Ile Thr Tyr Gly Thr  
 210 215 220  
 Ile Val Asp Gly Met Cys Lys Lys Gly Asp Thr Val Ser Ala Leu Asn  
 225 230 235 240  
 Leu Leu Arg Lys Met Glu Glu Val Ser His Ile Ile Pro Asn Val Val  
 245 250 255  
 Ile Tyr Ser Ala Ile Ile Asp Ser Leu Cys Lys Asp Gly Arg His Ser  
 260 265 270  
 Asp Ala Gln Asn Leu Phe Thr Glu Met Gln Glu Lys Gly Ile Phe Pro  
 275 280 285  
 Asp Leu Phe Thr Tyr Asn Ser Met Ile Val Gly Phe Cys Ser Ser Gly  
 290 295 300  
 Arg Trp Ser Asp Ala Glu Gln Leu Leu Gln Glu Met Leu Glu Arg Lys  
 305 310 315 320  
 Ile Ser Pro Asp Val Val Thr Tyr Asn Ala Leu Ile Asn Ala Phe Val  
 325 330 335  
 Lys Glu Gly Lys Phe Phe Glu Ala Glu Glu Leu Tyr Asp Glu Met Leu  
 340 345 350  
 Pro Arg Gly Ile Ile Pro Asn Thr Ile Thr Tyr Ser Ser Met Ile Asp  
 355 360 365  
 Gly Phe Cys Lys Gln Asn Arg Leu Asp Ala Ala Glu His Met Phe Tyr  
 370 375 380  
 Leu Met Ala Thr Lys Gly Cys Ser Pro Asn Leu Ile Thr Phe Asn Thr  
 385 390 395 400  
 Leu Ile Asp Gly Tyr Cys Gly Ala Lys Arg Ile Asp Asp Gly Met Glu

405	410	415
Leu Leu His Glu Met Thr Glu Thr Gly Leu Val Ala Asp Thr Thr Thr		
420	425	430
Tyr Asn Thr Leu Ile His Gly Phe Tyr Leu Val Gly Asp Leu Asn Ala		
435	440	445
Ala Leu Asp Leu Leu Gln Glu Met Ile Ser Ser Gly Leu Cys Pro Asp		
450	455	460
Ile Val Thr Cys Asp Thr Leu Leu Asp Gly Leu Cys Asp Asn Gly Lys		
465	470	475
Leu Lys Asp Ala Leu Glu Met Phe Lys Val Met Gln Lys Ser Lys Lys		
485	490	495
Asp Leu Asp Ala Ser His Pro Phe Asn Gly Val Glu Pro Asp Val Gln		
500	505	510
Thr Tyr Asn Ile Leu Ile Ser Gly Leu Ile Asn Glu Gly Lys Phe Leu		
515	520	525
Glu Ala Glu Glu Leu Tyr Glu Glu Met Pro His Arg Gly Ile Val Pro		
530	535	540
Asp Thr Ile Thr Tyr Ser Ser Met Ile Asp Gly Leu Cys Lys Gln Ser		
545	550	555
Arg Leu Asp Glu Ala Thr Gln Met Phe Asp Ser Met Gly Ser Lys Ser		
565	570	575
Phe Ser Pro Asn Val Val Thr Phe Thr Thr Leu Ile Asn Gly Tyr Cys		
580	585	590
Lys Ala Gly Arg Val Asp Asp Gly Leu Glu Leu Phe Cys Glu Met Gly		
595	600	605
Arg Arg Gly Ile Val Ala Asn Ala Ile Thr Tyr Ile Thr Leu Ile Cys		
610	615	620

Gly Phe Arg Lys Val Gly Asn Ile Asn Gly Ala Leu Asp Ile Phe Gln  
 625                      630                      635                      640  
 Glu Met Ile Ser Ser Gly Val Tyr Pro Asp Thr Ile Thr Ile Arg Asn  
                          645                      650                      655  
 Met Leu Thr Gly Leu Trp Ser Lys Glu Glu Leu Lys Arg Ala Val Ala  
                          660                      665                      670  
 Met Leu Glu Lys Leu Gln Met Ser Met Asp Leu Ser Phe Gly Gly  
                          675                      680                      685

<210> 30

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 30

acataaaaat cactagatac ttgacatgga ggc                      33

<210> 31

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 31

aagaggagga agatggcatc acagg                      25

<210> 32

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 32

tgaggtanag aggaactaaa aagggc 26

<210> 33

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 33

cagacaatag accgataaaa ggc 23

<210> 34

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 34

gattccttct tcttgcatct cag 23

<210> 35

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 35

atctcgtcct ttaccttctg tgg 23

<210> 36

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 36

gatccatgca ttgtcaagg 20

<210> 37

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 37

catttgtgta gctcatcta gg 22

<210> 38

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 38

gtcgggagag cagcccttgg tag 23

<210> 39

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic DNA

<400> 39

tcacgtata attcttcagc etc

23

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP02/04092

## A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl.<sup>7</sup> C07K14/415, C12N15/29, A01H1/00, A01H5/00, C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl.<sup>7</sup> C07K14/415, C12N15/29, A01H1/00, A01H5/00, C12Q1/68

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

SwissProt/PIR/GeneSeq, Genbank/EMBL/DBJ/Geneseq, WPI (DIALOG), BIOSIS (DIALOG)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 599042 A (Mitsubishi Corp.), 01 June, 1994 (01.06.94), & JP 06-141716 A & US 5644066 A	1-31
A	WO 92/05251 A (Inst. Nat. Rech. Agronomique), 02 April, 1992 (02.04.92), & EP 549726 A & JP 2001-145497 A	1-31
A	US 5973233 A (Zenco (No.4) Ltd.), 26 October, 1999 (26.10.99), & WO 97/02737 A	1-31
A	Hiroshi YAMAGISHI, "Distribution and allelism of restorer genes for Ogura cytoplasmic male sterility in wild and cultivated radishes.", Genes Genet.Syst., 1998, Vol.73, No.2, pages 79 to 83	1-31

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document relating to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search  
30 May, 2002 (30.05.02)Date of mailing of the international search report  
11 June, 2002 (11.06.02)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP02/04092

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Seiji MURAYAMA et al., Identification of RAPD and SCAR Markers Linked to a Restorer Gene for Ogura Cytoplasmic Male Sterility in Radish ( <i>Raphanus sativus</i> L.) by Bulked Segregant Analysis. Breeding Science, 1999, Vol.49, No.2, pages 115 to 121	1-31

## A. Classification of Field of the Invention (International Patent Classification (IPC))

Int. Cl. C07K14/415, C12N15/29, A01H1/00, A01H5/00, C12Q1/68

## B. Technical Field Searched

## Minimum Documentation of Search (International Patent Classification (IPC))

Int. Cl. C07K14/415, C12N15/29, A01H1/00, A01H5/00, C12Q1/68

Documentation included in the field of search in documentations other than minimum documentation

Electronic database used at International Search (Title of the database, language used in the search)

SwissProt/PIR/GeneSeq, Genbank/EMBL/DBJ/GeneSeq,  
WPI (DIALOG), BIOSIS (DIALOG)

## C. Documents considered to be relevant

Category of citation *	Title of Citation and indication of relevant section when a passage(s) is relevant	Related Claim Numbers
A	EP 599042 A(MITSUBISHI CORP.) 1994.06.01 & JP 06-141716 A & US 5644066 A	1-31
A	WO 92/05251 A(INST. NAT. RECH. AGRONOMIQUE) 1992.04.02 & EP 549726 A & JP 2001-145497 A	1-31
A	US 5973233 A(Zenco(No. 4)Limited) 1999.10.26 & WO 97/02737 A	1-31

☒ Further documents are listed in the continuation of Box C.☐ See separate paper regarding patent family.

## \* Categories of citations

[A] Document defining the general state of the art which is not considered to be of particularly relevance

[E] Earlier document but published on or after the international filing date

[L] Document which may throw doubts on priority claims(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

[O] Document referring to an oral disclosure, use, exhibition or other means

[P] Document published prior to the international filing date but later than the priority date claimed

[T] Later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle of theory underlying the invention

[X] Document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

[Y] Document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such document, such combinations being obvious to a person skilled in the art

[&amp;] Document member of the same patent family

Completion date of international search

30.05.02

Dispatch date of international  
search report

11.06.02

Title of International Search Organization and address thereof

Japan Patent Office (ISA/JP)

4-3, Kasumigaseki 3-chome, Chiyoda-ku, Tokyo

Zip code 100-8915

Examiner of IPO (employee with  
authority)

Koichi Niimi

Tel: 03-3581-1101 ext. 3448

4B 9162

## C (Continuation) Documents considered to be relevant

Category of citation *	Title of Citation and indication of relevant section when a passage(s) is relevant	Related Claim Numbers
A	Hiroshi YAMAGISHI. Distribution and allelism of restorer genes for Ogura cytoplasmic male sterility in wild and cultivated radishes. Genes Genet. Syst. 1998, Vol. 73, No. 2, p. 79-83	1-31
A	Seiji MURAYAMA et al. Identification of RAPD and SCAR Markers Linked to a Restorer Gene for Ogura Cytoplasmic Male Sterility in Radish ( <i>Raphanus sativus</i> L.) by Bulk Segregant Analysis. Breeding Science. 1999, Vol. 49, No. 2, p. 115-121	1-31